

Pd		25-SEP-2003.	
Xx			
Pf		17-MAR-2003; 2003MO-EP002735.	
Xx			
Xr		20-MAR-2002; 2002DE-01012892.	
Pr		(BADI ) BASF PLANT SCI GMBH.	
Px			
Pi	Kock M, Bauer J;		
Dx	WPI , 2003-603889/75.		
Dr	N-PsDB ; ADH89337 .		
Pt	Reducing expression of at least two target genes, useful e.g. for producing transgenic plants, using partly double-stranded interfering RNA.		
Ps	Disclosure; SEQ ID NO 113; 228bp; German.		
Cc	This invention describes a novel method for reducing the expression of at least two different endogenous target genes in a eukaryotic cell or organism by introducing an RNA molecule that is at least partly double stranded. The transcribed RNAs from at least two target genes have homology below 90% and the RNA molecule is formed as a single, self-complementary molecule. At least one of the double-stranded structures formed from individual sense sequences has an even number of repeats of 21 or 22 bp. The RNA molecule may include an intron-encoding sequence. At last two target genes are selected from different classes of storage protein genes, i.e. 2S albumen, 7S- or 11S/12S-globulin or zein-prolamine and at least one of the sense sequences is identical to storage protein sequences or genes in the homogentisate metabolic pathway or enzyme types, e.g. acetyl transferases, thioesterases, (de)branching enzymes or cellulases. The RNA of the invention, also related cassettes, expression systems, vectors and transgenic organisms are used for preparation of pharmaceuticals, in biotechnological processes and plant biotechnology, specifically in plants to improve protection against abiotic stress, to modify composition and/or content of fatty acids, lipid and oils, to modify carbohydrate composition, to alter colour or pigmentation, to reduce content of storage proteins, to increase resistance to pathogens, to inhibit stem break, to delay fruit ripening or aging, to induce male sterility, to reduce content of toxic or unwanted components, to modify lignification and/or lignin content, to modify the fibre component in foods or fibre quality in cotton, to reduce susceptibility to shock, to increase synthesis of Vitamin E, to reduce contents of nicotine, caffeine or theophylline and to increase methionine content, by reducing threonine biosynthesis. The method provides a rapid and efficient way of reducing gene expression, can inhibit more than one target gene, prevents development of multiple phenotypes (since the transcription rate is the same for all RNA sequences), significantly reducing the selection process required to produce an organism with effective suppression of all target genes), avoids problems of epigenetic gene silencing, does not require synthesis of individual RNA sequences CC and the method can be applied to plants with complex (polyploid) genomes. CC No interference between the individual RNA sequences occur. This sequence CC represents a protein encoded by a target gene used in the method of the invention.		
Sq	Sequence 307 AA:		
Oy	Query Match 30.8%; Score 440.5; DB 7; Length 307; Best Local Similarity 42.8%; Pred. No. 1.4e-29; Matches 125; Conservative 37; Mismatches 75; Indels 55; Gaps 16		
Db	13 PSCQQQPEGEVPLVYQQGQFP-----GGQQQPFGQGQRYPPSPSSQQLVLQLQAPPPQRLLP 67 :::                : 27 FGLERPMQQQPLRPDQTFFQQPLFSQQDQ----QLLFPQPSFSQQDP-----PFMQQQPP 78 ::            : 68 YLPQGSPPPCQP-----YPQPPGYSGPQGR-----SQQAQAQQQQQQQQQQQQQI--114 Db   79 FSQQQQLRLPQQPFSQQQQQLVLVP-QPPSSQQQQLVLPPQSSPFRQQQQQHQLVVQGIIP 137  115 -LQGITLQQLILPGMDVVLQGH-NIAH---ARSOTLVQOSTYLQLIQELCOHMLQIDPOS 167		

Db 138 VQPSLIGQLNAC-KVF:QDQCSPVAMQRLARSQMLQSSCHVMQDCCQQLPQLPQSS 196  
 QY 168 QCCALHNVTVAIIITHQCKKQCKQSSQVSFPQPIQQLPILQGSFRBPQKHPQ----- 219  
 Db 197 RVEAIRAIRIYSITL--QEQQQGVQSISQKQDPPQ--LDQCVSQPQDSQDQLGQQPQQ 251  
 QY 220 ---AGGS-VQPPQLPQPEIRNLALQTLPMACNVYIAPY--CTIAPGIGGT 265  
 Db 252 QQLAGQTFILQPHQIQMLEVMSIALRLIPTMCSNVVLYETTSVPEGV-GT 302  
 RESULT 8  
 ADG44134  
 ID ADG44134 standard; protein; 307 AA.  
 XX ADG44134;  
 AC  
 DT 26-FEB-2004 (first entry)  
 DE T. aestivum glutenin-1D1 protein.  
 XX  
 KW oil content; plant; storage protein; seed-specific promoter; 2S-albumin;  
 KW 7S-globulin; 11S-globulin; 12S-globulin; zein-prolamine; transgenic;  
 KW oil production; fct production; free fatty acid production; food;  
 KW animal feed; pharmaceutical; fine chemical production; glutenin.  
 XX  
 OS Triticum aestivum.  
 XX MO2003077643-A2.  
 XX  
 PN 25-SEP-2003.  
 XX  
 PF 17-MAR-2003; 2003WO-EP002733.  
 XX  
 PF 20-MAR-2002; 2002DE-01012893.  
 XX  
 PA (BADI ) BASF PLANT SCI GMBH.  
 XX  
 PI Bauer J;  
 XX  
 DR WPI; 2004-011465/01.  
 DR N-PSDB; ADG44133.  
 PT  
 PT Increasing total oil content of plants, useful e.g. as foods or animal  
 PT feeds, by reducing amount of storage proteins, particularly with double-  
 PT stranded interfering RNA.  
 XX  
 XX Claim 4; SEQ ID NO 174; 253bp; German.  
 XX  
 XX This invention describes a novel method for increasing the total oil  
 XX content of a plant by reducing the amount of at least one storage protein  
 XX in the plant (or its tissue, organs, parts or cells) and selecting plants  
 XX that have higher total oil content than starting plants. The storage  
 XX protein is suppressed by introducing antisense RNA, optionally combined  
 XX with a ribozyme, sense RNA that induces co-suppression, DNA-binding  
 XX factors directed against storage protein genes, viral sequences that  
 XX degrade storage protein RNA, constructs that induce homologous  
 XX recombination of endogenous storage protein genes or mutations into  
 XX storage protein genes. Most preferably a plant cell is stably transfected  
 XX with a recombinant expression construct, then regenerated to plants that  
 XX express the incorporated sequence. The expression constructs particularly  
 XX contain a seed-specific promoter and they are introduced into plants by  
 XX standard methods, e.g. via Agrobacterium. The preferred storage proteins  
 XX of the invention are 2S-albumens, 7S or 11S/12S-globulins or zein-  
 XX prolamines. Transgenic organisms produced by the new method are used for  
 XX production of oils, fats, free fatty acids or their derivatives, useful  
 XX as foods, animal feeds, pharmaceuticals and fine chemicals. This sequence  
 XX represents a storage protein used to illustrate the method of the  
 XX invention.  
 XX  
 XX Sequence 307 AA;  
 XX







DR N-PSDB; ABL15798.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 PT interactions.  
 PT  
 XX  
 PS Disclosure; SEQ ID NO 41877; 21bp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signaling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA  
 CC sequences (AB101840-AB16175) and the encoded proteins (AB57737-  
 CC AB572072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 CC  
 XX  
 SQ Sequence 1798 AA;  
 Query Match 20.2%; Score 288.5; DB 4; Length 1798;  
 Best Local Similarity 29.5%; Pred. No. 1.2e-15;  
 Matches 102; Conservative 18; Mismatches 91; Indels 135; Gaps 12;  
 QY 4 PVVQLQPNPSQQPQEQVPLVQQQPFPGQQQF---PQQPYPPQPPSPQQPYQLQLP 60  
 DB 150 PAVQMTTPPPSPQNAQQQHLRQQQQQCNQCNQMMGSPQPPQPSQMSPPQQQ--QMQP 207  
 QY 61 PPOP-----RLPYQ-----PQSFPQPFY----- 80  
 DB 208 PQPQVHQQNMQLQQQQQLAQQQQQQQQSPHISFQSPQISQTPPMQAKLHQVVPQ 267  
 QY 81 PQQPQYSQ-----PQQPIS-----QQQAQQQQQQQQQ----- 109  
 DB 268 ATPQCSFSQCKPIDPDPVQAVLSRSALSSNQDSLIMRQQQLKQQQMQQQQQQMAPQ 327  
 QY 110 -QQQQQLLQQLLPCQMVVLAQHNIAHBSQVLAQSYQLQLCCQHLMQPESQ 168  
 DB 328 PQQQQMAQPPQQQ---QQQPQQQTPSPRQSPQQQPTPLQ-----QQPNQON 375  
 QY 169 CQAHHVVAHLLHQQKQKQSPSSQVSPQPPQYPLGSGSFRPSQ----- 216  
 DB 376 AQT-----QQQQQQQQQQQQQQQQQQQVLTQQQPQPPQQQQVTRHVVNT 425  
 QY 217 -----NPAQGSVPPQQLPQREIRNL 238  
 DB 426 STRAQGQIIGSHMSLALQKQQLLHVQQQAQQQPPQQQQQITTVQQL 471  
 RESULT 11  
 ID AAY96255 standard; protein; 1162 AA.  
 AC AAY96255;  
 XX  
 DT 12-SEP-2003 (revised)  
 DT 11-SEP-2000 (first entry)  
 XX  
 DE Kaposi's sarcoma-associated herpesvirus LANA.  
 XX  
 KM Kaposi's sarcoma-associated herpesvirus; KSHV; rhadino virus;  
 KM latency-associated nuclear antigen; LANA; gamma-2 herpes virus;  
 KM Human herpes virus 8; HHV8; rhadino virus cis-acting element; RYCAE;  
 KM Kaposi's sarcoma; Primary effusion lymphoma; PEL;  
 KM human immunodeficiency virus; HIV; multicentric Castlemans disease.  
 XX  
 OS Human herpesvirus 8.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 14..17  
 FT Domain /note="nuclear localisation signal, NLS"  
 FT Domain 64..70

FT  
 FT Region /note="nuclear localisation signal, NLS"  
 FT 320..429  
 FT /note="acidic repeat region"  
 FT 430..549  
 FT Region /note="Gln, Glu, Pro-rich region"  
 FT 550..589  
 FT /note="Gln, Glu, Pro, Arg-rich region"  
 FT 590..759  
 FT Region /note="Gln, Glu, Asp-rich region"  
 FT 760..840  
 FT /note="Gln, Glu-rich region"  
 FT  
 FN WO200029626-A1.  
 PD 25-MAY-2000.  
 PP 19-NOV-1999; 99WO-US027508.  
 PR 19-NOV-1998; 98US-00109422.  
 PR 21-APR-1999; 99US-00298568.  
 XX  
 PA (KIEF/) KIEF E D.  
 PA (BALL/) BALLESTAS M E.  
 PA (KAYE/) KAYE K M.  
 PI Kieff ED, Ballestas ME, Kaye KM;  
 XX  
 DR WPI; 2000-387829/33.  
 DR N-PSDB; AAA30290.  
 PT  
 PT Treating or preventing a disease associated with rhadino virus infection  
 PT in a mammal which includes Kaposi's Sarcoma and Primary Effusion  
 PT Lymphoma.  
 PS  
 PS Disclosure; Fig 7; 70P; English.  
 XX  
 CC The present sequence is the Kaposi's sarcoma-associated herpesvirus,  
 CC (KSHV) latency-associated nuclear antigen (LANA). KSHV is also known as  
 CC Human Herpes Virus 8 (HHV8) and belongs to the rhadino virus, or gamma-2  
 CC herpes virus class. The LANA protein is necessary for the efficient  
 CC persistence of rhadino virus DNA in mammalian cells. Persistent rhadino  
 CC virus infection is implicated in a variety of diseases e.g. Kaposi's  
 CC Sarcoma (KS), Primary Effusion Lymphoma (PEL) and multicentric  
 CC Castlemans disease. In addition, KS is a common malignancy in HIV  
 CC patients. KSHV persists in host cells in a latent form. One of the few  
 CC genes expressed from the latent viral DNA is LANA. LANA associates with  
 CC both human chromosomes and with the rhadino virus cis-acting element  
 CC (RYCAE), thereby providing a tethering function: the KSHV DNA episome is  
 CC "tied" to the host chromosomes. This allows the viral DNA to persist in  
 CC the host cell. The present sequence may be used to screen and identify  
 CC molecules that inhibit LANA interaction with RYCAE, thereby interfering  
 CC with the latency cycle of this virus. Potential antiviral treatments for  
 CC the above mentioned diseases may therefore be based on LANA deregulation.  
 CC (Updated on 12-SEP-2003 to standardise OS field)  
 CC  
 XX  
 SQ Sequence 1162 AA;  
 Query Match 19.1%; Score 273; DB 3; Length 1162;  
 Best Local Similarity 38.4%; Pred. No. 1.5e-14;  
 Matches 94; Conservative 19; Mismatches 100; Indels 32; Gaps 10;  
 QY 4 PVVQLQPNPSQQPQEQVPLVQ---QQQFPGQ---QQPFPQPYPPQPPSPQQPYQLQLP 59  
 DB 495 PLOEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEP 553  
 QY 60 P---PQPRLPY---PQ---PQSFPQPPYPPQ---QPPQSPQPPQISQQAQQQQQQQQQQ 109  
 DB 554 PQQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPP 613  
 QY 110 QQQQQLLQQLLPCQMVVLAQHNIAHBSQVLAQSYQLQLCCQHLMQPESQ 169  
 DB 614 EQQQDEQQQDEQQQ---QDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDE 663



QY 1.70 GAHHNVHAIILHQQKQKQSSQVSPQPLQGYPLGGGFRPSQGNFQAQSVQPOL 229  
 DB 664 EQQOD-----EQQDEQDDEQDDEQDDEQDDEQDDEQDDEQDDEQDDEQD 716  
 QY 230 PQFEE 234  
 DB 717 EQQOD 721

RESULT 12  
 AAY58500 standard; protein; 1162 AA.  
 AAY58500;  
 06-AUG-2003 (revised)  
 10-APR-2000 (first entry)  
 HHV8 ORF 73 protein, SEQ ID NO:21.  
 HHV8, detection; diagnosis; Kaposi's sarcoma; AIDS; immunogen; antigen.  
 Human herpesvirus 8.  
 Key Location/Qualifiers  
 Misc-difference 96  
 /label= unknown  
 W09961909-A2.  
 02-DEC-1999.  
 26-MAY-1999; 99MO-US011407.  
 26-MAY-1999; 98US-0086695P.  
 (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 Pau C;  
 WPI; 2000-097142/08.  
 New methods and compositions for the detection of human herpesvirus.  
 Claim 2; Page 59-62; 68pp; English.  
 Sequences AAY58480-Y58532 represent immunogenic polypeptides derived from human herpes virus type 8 (HHV8, a gammaherpesvirus). HHV8 plays an important role in the pathogenesis of AIDS-related Kaposi's sarcoma. The invention relates to a novel method of detecting the presence of human herpesvirus 8 in a biological sample using peptides representative of dominant antigenic regions of HHV8. The method comprises contacting one or more isolated, immunogenic HHV8 peptides with an antibody-containing biological sample, and detecting the formation of a complex between the peptide and the antibody. The presence of a peptide-antibody complex indicates the presence of human herpesvirus 8. The detection of HHV8 infection can be used to diagnose AIDS-associated Kaposi's sarcoma. The HHV8-specific antibodies are useful therapeutically when for the passive immunisation of a human against HHV8 infection, thereby reducing HHV8 related disease. The detection assays are highly specific, sensitive and accurate. Early detection and treatment of Kaposi's sarcoma could diminish the severity of symptoms related to AIDS and the sensitive techniques could reduce erroneous characterisations of skin disorders. Previous assays for HHV8 antibodies such as immunofluorescence assays, immunoblots and enzyme immunoassays lack the sensitivity and accuracy needed for reliable diagnosis of Kaposi's sarcoma. Further advantages of the assays are that reproducible results are obtained and the method is suitable for rapid throughput and screening of samples economically.  
 (Updated on 06-AUG-2003 to correct OS field.)  
 Sequence 1162 AA;  
 Query Match 19.1%; Score 273; DB 3; Length 1162;

Best Local Similarity 38.4%; Pred. No. 1,5e-14;  
 Matches 94; Conservative 19; Mismatches 100; Indels 32; Gaps 10;  
 QY 4 PVPQLPQNPNSQKQDQGVPLVQ--QQQFPQD--QPPPPQPPPPQPPPPQPPVQLQ 59  
 DB 495 PLOEPQD 553  
 QY 60 P---PQPPPLPY---PQ---PQSPFQDQPPPP--QPPSPQPPPLSQDPAQDQDQDQDQ 109  
 DB 554 PQQREPPQDQ 613  
 QY 110 QQQQILQQLLQQLLPCMDVVLQCHNIAHNSQVLYQSTYGLQELCCOHLMQIPESQC 169  
 DB 614 EQQDEQDQ 663  
 QY 1.70 GAHHNVHAIILHQQKQKQSSQVSPQPLQGYPLGGGFRPSQGNFQAQSVQPOL 229  
 DB 664 EQQOD-----EQQDEQDDEQDDEQDDEQDDEQDDEQDDEQDDEQDDEQ 716  
 QY 230 PQFEE 234  
 DB 717 EQQOD 721

RESULT 13  
 AAB62331  
 ID AAB62331 standard; protein; 1162 AA.  
 AAB62331;  
 06-AUG-2003 (revised)  
 29-JUN-2001 (first entry)  
 Amino acid sequence of KSHV tethering protein LANA.  
 Histone H1; tethering protein; LANA; gene therapy; multiple sclerosis; Parkinson's disease; Huntington disease; diabetes; human herpesvirus 8; KSHV; latency-associated nuclear antigen; LANA.  
 Human herpesvirus 8.  
 W0200125484-A2.  
 12-APR-2001.  
 29-SEP-2000; 2000MO-US026908.  
 01-OCT-1999; 99US-00410399.  
 (UNMI ) UNIV MICHIGAN.  
 Robertson ES, Cotter MA;  
 WPI; 2001-281736/29.  
 N-PSDB; AAF82901.  
 A composition for use in gene therapy comprises an expression vector that includes a nucleic acid sequence encoding a nucleic acid binding protein.  
 Disclosure; Fig 9B; 60pp; English.  
 The invention provides a composition comprising nucleic acid, histone H1 protein and expression vector operatively encoding a protein suitable for tethering the nucleic acid to the histone H1 protein, where the tethering protein is LANA. The composition is useful in aiding the retention of the viral DNA in the host cell. The viral vector encodes a protein suitable for tethering DNA to histone H1. Methods for screening for compounds which are agonistic or antagonistic for the tethering of viral proteins to histone H1 and DNA binding sites are useful for developing the method of viral transfer. The composition has applications to gene therapy, including the treatment of multiple sclerosis, Parkinson's disease, Huntington disease and diabetes. The present sequence represents the amino acid sequence of the Kaposi's sarcoma



CC associated herpesvirus (human herpesvirus 8) latency-associated nuclear  
CC antigen (LANA), which acts as a tethering protein. (Updated on 06-AUG-  
CC 2003 to correct OS field.)  
XX  
XX  
SQ Sequence 1162 AA;

Query Match 19.1%; Score 273; DB 4; Length 1162;  
Best Local Similarity 38.4%; Pred. No. 1.5e-14;  
Matches 94; Conservative 19; Mismatches 100; Indels 32; Gaps 10;

QY 4 PVPOLQPNPSQQQPGQVPLVQ--QQQPFQQQ--QQFPQQPYPPQPPPSQQPYLQLQ 59  
DB 495 PLOEPQQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQ 553  
QY 60 P--FPQPLPY--PQ--PQSFPPQPYPP--QPQYSQPQQPISQQQAQQQQQQQQQQ 109  
DB 554 PQQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPP 613  
QY 110 QQQQILQQILQQQLPCMDVYLQGHNTAHARSQVLTQSTYQLLQELCCGHLMQIPBSQC 169  
DB 614 EQQDDEQQDEQQ--QDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQD 663  
QY 170 QAHNVVHAITLHQGQKQQQPPSSQVSFPQPLQGYPLGGGSRPPSQQNFQAGSVQPPQL 229  
DB 664 EQQD-----EQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQD 716  
QY 230 PQQEE 234  
DB 717 EQQD 721

## RESULT 14

ABB05621  
ID ABB05621 standard; protein; 1162 AA.

XX  
AC ABB05621;

XX  
DT 25-APR-2002 (first entry)

XX  
DE Kaposi's sarcoma-associated herpesvirus LANA protein.

XX  
KW Kaposi's sarcoma-associated herpesvirus; KSHV; LANA; RVCAR; PEL;  
KW KSHV terminal repeat; rhadino virus cis acting element; episome;  
KW primary effusion lymphoma; latency-associated nuclear antigen;  
KW gene therapy; gene transfer.

XX  
OS Human herpesvirus 8.

XX  
PN US6322792-31.

XX  
PD 27-NOV-2001.

XX  
PE 21-APR-1999; 99US-00298568.

XX  
PR 19-NOV-1998; 98US-0109422P.

XX  
PA (KIEF/) KIEFF E D.

XX  
PI Kieff ED, Ballestas ME, Kaye KM;

XX  
DR WPI: 2002-153769/20.

XX  
N-PSDB; ABA93487.

XX  
PT System for episomal retention of plasmids in mammalian cells, useful in  
gene therapy, comprises rhadinoviral LANA and RVCAR sequences.

XX  
PS Disclosure; Fig 7; 27pp; English.

XX  
CC The present invention describes a system (A) for maintaining a plasmid as  
an episome in mammalian cells, comprising the rhadinoviral sequence LANA  
CC (latency-associated nuclear antigen) of 3469 base pairs (see ABA93487,  
CC S1) expressed in the cell, and the rhadinoviral sequence RVCAR  
CC (rhadinoviral cis-acting element) of 801 base pairs (see ABA93488, S2).

CC present in the plasmid. Also describes is a method for maintaining a  
CC closed circular DNA in a cell by expressing (S1) in the cells and having  
CC (S2) as a cis-acting and maintenance sequence in the DNA. (A) is  
CC particularly used in gene therapy (or other gene transfer applications)  
CC that uses mammalian cells in which LANA is expressed. (A) improves  
CC persistence of gene therapy vectors in cells. The present sequence  
CC represents Kaposi's sarcoma-associated herpesvirus (KSHV, also called  
CC human herpesvirus 8) LANA protein, which is used in the exemplification  
CC of the present invention

Query Match 19.1%; Score 273; DB 5; Length 1162;  
Best Local Similarity 38.4%; Pred. No. 1.5e-14;  
Matches 94; Conservative 19; Mismatches 100; Indels 32; Gaps 10;

QY 4 PVPOLQPNPSQQQPGQVPLVQ--QQQPFQQQ--QQFPQQPYPPQPPPSQQPYLQLQ 59  
DB 495 PLOEPQQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQEPQQ 553  
QY 60 P--FPQPLPY--PQ--PQSFPPQPYPP--QPQYSQPQQPISQQQAQQQQQQQQQQ 109  
DB 554 PQQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPPQREPP 613  
QY 110 QQQQILQQILQQQLPCMDVYLQGHNTAHARSQVLTQSTYQLLQELCCGHLMQIPBSQC 169  
DB 614 EQQDDEQQDEQQ--QDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQD 663  
QY 170 QAHNVVHAITLHQGQKQQQPPSSQVSFPQPLQGYPLGGGSRPPSQQNFQAGSVQPPQL 229  
DB 664 EQQD-----EQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQDDEQQD 716  
QY 230 PQQEE 234  
DB 717 EQQD 721

## RESULT 15

ADJ65096  
ID ADJ65096 standard; protein; 1162 AA.

XX  
AC ADJ65096;

XX  
DT 20-MAY-2004 (first entry)

XX  
DE HHV8 latency-associated nuclear antigen, LANA.

XX  
KW HHV8; latency-associated nuclear antigen; LANA;  
KW primary effusion lymphoma; PEL; virucide; gene therapy; KSHV;  
KW Kaposi's sarcoma-associated herpesvirus; episome; RVCAR;  
KW rhadinovirus cis-acting element; rhadinovirus infection.

XX  
OS Human herpesvirus 8.

XX  
PN US2004037847-A1.

XX  
PD 26-FEB-2004.

XX  
PE 28-JUN-2001; 2001US-00894273.

XX  
PR 19-NOV-1998; 98US-0109422P.

XX  
PR 21-APR-1999; 99US-00298568.

XX  
PA (KIEF/) KIEFF E D.

XX  
PA (BALU/) BALLESTAS M E.

XX  
PA (KAYE/) KAYE K M.

XX  
PI Kieff ED, Ballestas ME, Kaye KM;

XX  
CC Assays for compounds that modulate rhadino virus LANA action in trans on







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 17:06:20 ; Search time 19 Seconds  
(without alignments)  
928,452 Million cell updates/sec

Title: US-10-089-700-3-R65  
Perfect score: 1431  
Sequence: 1 VAVPVQQLPQNPQQQPE.....CNVYAPYCTIAFCIFGTN 266

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 segs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/prodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/prodata/1/iaa/PCMTUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/1/iaa/backfillset.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	464	32.4	369	2	US-08-991-300-2
2	273	19.1	1162	2	US-08-728-323A-2
3	273	19.1	1162	3	US-09-298-568-2
4	273	19.1	1162	4	US-09-410-399-2
5	273	19.1	1162	4	US-09-894-273-2
6	249	17.4	788	2	US-08-918-914-4
7	234	16.4	256	4	US-09-248-796A-21251
8	232	16.2	498	4	US-09-270-767-45042
9	217	15.2	579	4	US-09-668-119-3
10	217	15.2	2074	4	US-09-481-356C-9
11	202	14.1	2023	4	US-09-481-356C-8
12	202	14.1	2124	4	US-09-538-092-1377
13	198.5	13.9	505	4	US-09-248-796A-19253
14	197.5	13.8	663	4	US-09-270-767-41220
15	197.5	13.8	1591	4	US-09-270-767-45698
16	197.5	13.8	2441	1	US-08-194-468-2
17	197.5	13.8	2441	3	US-08-961-739-2
18	197.5	13.8	2441	3	US-09-514-247A-8
19	197.5	13.8	2441	4	US-09-686-316-2
20	196.5	13.7	379	4	US-09-248-796A-23759
21	196.5	13.7	2442	3	US-09-514-247A-10
22	196.5	13.7	2442	4	US-09-538-092-1370
23	196	13.7	216	4	US-09-248-796A-21017
24	195.5	13.7	729	3	US-09-625-188-20
25	188	13.1	295	4	US-09-248-796A-20004
26	186.5	13.0	320	4	US-09-248-796A-24758
27	185	12.9	169	4	US-09-248-796A-28087

28	184.5	12.9	316	4	US-09-270-767-42663	Sequence 42663, A
29	184.5	12.9	332	4	US-09-248-796A-21649	Sequence 21649, A
30	184	12.9	519	4	US-09-248-796A-19263	Sequence 19263, A
31	178	12.4	383	4	US-09-248-796A-23236	Sequence 23236, A
32	178	12.4	408	4	US-09-248-796A-14480	Sequence 14480, A
33	178	12.4	1319	4	US-09-538-092-1291	Sequence 1291, Ap
34	177	12.4	261	4	US-09-602-565-34	Sequence 34, Appl
35	176.5	12.3	684	4	US-09-823-240A-9	Sequence 15319, A
36	176	12.3	618	4	US-09-248-796A-15319	Sequence 19232, A
37	176	12.3	657	4	US-09-248-796A-19232	Sequence 33, Appl
38	176	12.3	848	4	US-09-538-092-93	Sequence 18720, A
39	174.5	12.2	382	4	US-09-248-796A-18720	Sequence 3, Appl
40	173	12.1	1507	4	US-09-914-259-37	Sequence 27827, A
41	171.5	12.0	311	4	US-09-248-796A-27827	Sequence 2, Appl
42	171.5	12.0	903	4	US-08-853-310-2	Sequence 20699, A
43	171	11.9	675	4	US-09-248-796A-20699	Sequence 2, Appl
44	168.5	11.8	667	2	US-08-718-661-2	Sequence 6, Appl
45	168.5	11.8	107	4	US-09-668-119-6	

## ALIGNMENTS

RESULT 1  
US-08-991-300-2  
Sequence 2, Application US/08991300  
Patent No. 5973225  
GENERAL INFORMATION:  
APPLICANT: D'OVIDIO, RENATO  
APPLICANT: BORCEDDU, ENRICO  
APPLICANT: MERCHITELLI, CINZIA  
APPLICANT: CARDELLI, LUISA ERCOLI  
TITLE OF INVENTION: ISOLATION AND CHARACTERIZATION OF A GENE  
TITLE OF INVENTION: ENCODING A LOW MOLECULAR WEIGHT GLUTENIN  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESS: OBLON, SPIVAK, MCCLELLAND, NAIER & NEUSTADT,  
ATTORNEYS  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/991,300  
FILING DATE: 16-DEC-1997  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: IT MI 96/A 002663  
FILING DATE: 19-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 2264-0201-0X  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 369 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-991-300-2  
Query Match 32.4%, Score 464, DB 2, Length 369;  
Best Local Similarity 40.2%, Pred. No. 9, 2e-35;







```

US-09-410-399-2
; Sequence 2, Application US/09410399
; Patent No. 6482587
; GENERAL INFORMATION:
; APPLICANT: Robertson, Erle S.
; TITLE OF INVENTION: Methods to Inhibit or Enhance the Binding of Viral DNA
; TITLE OF INVENTION: to Genomic Host DNA
; FILE REFERENCE: UM-03778
; CURRENT APPLICATION NUMBER: US/09/410,399
; CURRENT FILING DATE: 1999-10-01
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1162
; TYPE: prt
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-410-399-2

```

[illegible]

```

RESULT 5
US-09-894-273-2
; Sequence 2, Application US/09894273
; Patent No. 6756203
; GENERAL INFORMATION:
; APPLICANT: Kieff, Elliott D.
; APPLICANT: Balleskas, Mary E.
; TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO
; TITLE OF INVENTION: VIRUS DNA TO MEDIANE EFFICIENT EPISOME PERSISTENCE
; FILE REFERENCE: 16412-10001R
; CURRENT APPLICATION NUMBER: US/09/894,273
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/109,422
; PRIOR FILING DATE: 1998-11-19
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ. ID NO. 2
; LENGTH: 1162
; TYPE: prt
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-894-273-2

```

```

Query Match      19.1%; Score 273; DB 4; Length 1162;
Best Local Similarity 38.4%; Freq. No. 1,2e-16;
Matches 94; Conservative 19; Mismatches 100; Indels 33; Gaps 10.

QY      4 PVPQLPQNMSSQQQPOEQLVPLVQ--QQQSPGGG--QQFPQQQYPCQPPQSPQAPYLQIQ 59
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |

```

[illegible]

```

1      RESULT 6
2      US-08-918-914-4
3      ; Sequence 4, Application US/08918914
4      ; Patent No. 5876963
5      ; GENERAL INFORMATION:
6      ; APPLICANT: Mitchell, Peter
7      ; APPLICANT: Hutchinson, Nancy
8      ; APPLICANT: Lawton, Michael
9      ; APPLICANT: Magna, Holly
10     ; APPLICANT: Yocum, Sue
11     ; APPLICANT: Murry, Lynn E.
12     ; TITLE OF INVENTION: HUMAN NUCLEOTIDE PYROPHOSPHORYLASE
13     ; NUMBER OF SEQUENCES: 4
14     ; CORRESPONDENCE ADDRESS:
15     ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
16     ; STREET: 3174 Porter Dr.
17     ; CITY: Palo Alto
18     ; STATE: CA
19     ; COUNTRY: USA
20     ; ZIP: 94304
21     ; COMPUTER READABLE FORM:
22     ; MEDIUM TYPE: Diskette
23     ; COMPUTER: IBM Compatible
24     ; OPERATING SYSTEM: DOS
25     ; SOFTWARE: Pasteo for Windows Version 2.0
26     ; CURRENT APPLICATION DATA:
27     ; APPLICATION NUMBER: US/08/918,914
28     ; FILING DATE: Filed Herewith
29     ; PRIOR APPLICATION DATA:
30     ; APPLICATION NUMBER:
31     ; FILING DATE:
32     ; ATTORNEY/AGENT INFORMATION:
33     ; NAME: Billings, Lucy J.
34     ; REGISTRATION NUMBER: 36,749
35     ; REFERENCE/DOCKET NUMBER: PF-0369
36     ; TELECOMMUNICATION INFORMATION:
37     ; TELEPHONE: 415-855-0555
38     ; TELEFAX: 415-845-4166
39     ; TELEX:
40     ; INFORMATION FOR SEQ ID NO: 4:
41     ; SEQUENCE CHARACTERISTICS:
42     ; LENGTH: 788 amino acids
43     ; TYPE: amino acid
44     ; STRANDEDNESS: single
45     ; TOPOLOGY: linear
46     ; IMMEDIATE SOURCE:
47     ; LIBRARY: Genbank
48     ; CLONE: 1070094
49     ;
50     ; US-08-918-914-4

```

Query Match	17.4%	Score 249;	DB 2;	Length 788;
Best Local Similarity	31.7%	Pred NO 1.2e-14;		
Matches 91; Conservative	22;	Mismatches 96;	Indels 78;	Gaps 11.



QY 2 RVFVPOLO-----PQNFSCQOPQOEQVPLVQOQFPQOQOQFPQOQVPPQOQPPPSQOY 55  
 DB 201 RVFAPPAQOVARAPRAPEVPSAQOQOQ-----QREOQOQBREHQAQLQOQOQOQOQ 255  
 QY 56 LQLOFPQRLPYPPQOQFPQOQFPQOQFPQOQFPQOQFPQOQFPQOQFPQOQFPQOQFPQOQ 115  
 DB 256 QOQOQFPQOQ---FQPOQFPQOQFPQOQ---SFGTHLHLQOQOQOQOQOQOQOQOQ 311  
 QY 116 QOQLOQOQILPCMDVVLQOQNIHAHSQVLAQSTVQLOLQELCCOHLMOIPROSCQAIHNV 175  
 DB 312 QONPQOQ-----PQOTLQFGSQOIQLOQS-----GVPFPOQH----- 342  
 QY 176 VHAIIHQQOQOQOQSSQVS-----FQOPLQOY-----PLQGSFR 212  
 DB 343 -----PQOQOQOQOQBELERSPLQOHAQLYOQKMSQIRENNGHPRAPKADPCGSGFCA 396  
 QY 213 PSQGNPAQSGVPPQOLPQFEIRNLALQTLPAACNVYIAPYCTIAP 259  
 DB 397 PVPQAPQOE---RFTPPVLAIVINTATQ---PPLPQVPTRYRPA 438

## RESULT 7

US-09-248-796A-21251

Sequence 21251, Application US/09248796A

Patent No. 6747137

GENERAL INFORMATION:

APPLICANT: Keith Weinstein et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN

FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.132

CURRENT APPLICATION NUMBER: US/09/248, 796A

CURRENT FILING DATE: 1999-02-12

PRIOR APPLICATION NUMBER: US 60/074,725

PRIOR FILING DATE: 1998-02-13

PRIOR APPLICATION NUMBER: US 60/096,409

PRIOR FILING DATE: 1998-08-13

NUMBER OF SEQ ID NOS: 28208

SEQ ID NO 21251

LENGTH: 256

TYPE: PRT

ORGANISM: Candida albicans

FEATURE:

NAME/KEY: UNSURE

LOCATION: (250)

OTHER INFORMATION: Identity of amino acid sequences at the above locations are unkno

US-09-248-796A-21251

Query Match 15.4%; Score 234; DB 4; Length 256;  
 Best Local Similarity 36.8%; Pred. No. 7.2e-14;

Matches 74; Conservative 11; Mismatches 66; Indels 50; Gaps 8;

QY 31 PQOQOQFPQOQVPPQOQFPQOQVPLQOFPQOQVPPQOQFPQOQVPPQOQVPPQOQVPPQOQ 90  
 DB 75 PQOQOQFPQOQVPPQOQFPQOQVPLQOFPQOQVPPQOQFPQOQVPPQOQVPPQOQVPPQOQ 119  
 QY 91 QOQFISQOQ 150  
 DB 120 QOQFISQOQ 160  
 QY 151 LQOELCCOHLMOIPROSCQAIHNVVHAIIHQQOQKQOQOQSSQVSFQOPLQOQVPLQOQ 210  
 DB 161 PPOQ---QHYKQOQ 206  
 QY 211 FRSQOQNPQOQSGVPPQOQFPQOQ 231  
 DB 207 QPPPOQ---QLYGRSQOQOQOQ 225

## RESULT 8

US-09-270-767-45042

Sequence 45042, Application US/09270767

Patent No. 6703491

GENERAL INFORMATION:

APPLICANT: Homburger et al.  
 TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster  
 FILE REFERENCE: File Reference: 7326-094  
 CURRENT APPLICATION NUMBER: US/09/270,767  
 CURRENT FILING DATE: 1999-03-17  
 NUMBER OF SEQ ID NOS: 62517  
 SOFTWARE: Patent In Ver. 2.0  
 SEQ ID NO 45042  
 LENGTH: 498  
 TYPE: PRT  
 ORGANISM: Drosophila melanogaster  
 FEATURE:  
 OTHER INFORMATION: Xaa means any amino acid  
 US-09-270-767-45042

Query Match

15.2%; Score 232; DB 4; Length 498;  
 Best Local Similarity 32.1%; Pred. No. 2.4e-13;

Matches 99; Conservative 20; Mismatches 101; Indels 88; Gaps 15;

QY 6 POLQPNP-SQOQPOQOQVPLVQOQOQFPQOQOQFPQOQVPPQOQFPQOQVPLQOQ 57  
 DB 181 PQLQTAGPQOQ 235  
 QY 58 ---LQFPQOQVPPQOQFPQOQFPQOQFPQOQFPQOQFPQOQFPQOQFPQOQFPQOQ 105  
 DB 236 MOELQOFGQDFQLABSNTS-PPQO---OQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ 291  
 QY 106 OQ 151  
 DB 292 QOQOQOQVPLVQOQVPLQOQALSGPHVPPQOQOQALPQOQVPHMQKQO---QO 348  
 QY 152 LQELCCOHLMOIPROSCQAIHNVVHAIIHQQOQKQOQOQSSQV---SFGQPLQOQVPLQO 208  
 DB 349 LVETQOHV---OKQ-----HOSQPOVOQPPQOQLOQPSQOQVPLPY--- 386  
 QY 209 GSFPSQONPQO-----QSVQOQ 257  
 DB 387 HTMPPOQSPVNVTVSVPLLEQPPQPMVVOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ 440  
 QY 258 APEGIFGT 265  
 DB 441 TPTGIAS 448

## RESULT 9

US-09-668-119-3

Sequence 3, Application US/09668119

Patent No. 6768003

GENERAL INFORMATION:

APPLICANT: Solomon, William B

TITLE OF INVENTION: Transcriptional adaptor protein

FILE REFERENCE: 011.00250

CURRENT APPLICATION NUMBER: US/09/668,119

CURRENT FILING DATE: 2000-09-22

NUMBER OF SEQ ID NOS: 8

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 3

LENGTH: 579

TYPE: PRT

ORGANISM: Homo sapiens

US-09-668-119-3

Query Match 15.2%; Score 217.5; DB 4; Length 579;  
 Best Local Similarity 33.9%; Pred. No. 6.2e-12;  
 Matches 93; Conservative 15; Mismatches 87; Indels 79; Gaps 12;

QY 4 PVPOLQPNP-SQOQPOQOQVPLVQOQOQFPQOQOQFPQOQVPPQOQFPQOQVPLQOQ 61  
 DB 123 PQTQQLQOQVALQOQ 176  
 QY 62 ---PQFRLPYPPQOQFPQOQFPQOQFPQOQFPQOQFPQOQFPQOQFPQOQFPQOQFPQOQ 115



```

Query Match Similarity 15.2%; Score 217; DB 4; Length 2074;
Best Local Similarity 30.6%; Pred. No. 3.2e-11;
Matches 87; Conservative 15; Mismatches 90; Indels 92; Gaps 10;

QY      7  QLOPNECQQQOEVEVLVQQQQQFFGQQ--QQFPQQPY-----PQQPFPSSQQPY 55
Db      1807  QQQPFPVPGGQRRLQQ--LQQSQGMLGGSSVHOMTPSSSYGLQTSLSPPSLQGYTIVSH 1864

QY      56  LQLOPFPQRLPYEPQSQFPQQPYPQPOPOYSQ-----89
Db      1865  VGLQQHTBPADPTRRLQGRPSGVYH--QQAFTYGHGLTGTGRFSSHQTLQOTPEMGMTPLS 1923

QY      90  PQQPISQQQAQQQQQQQQQQQQQQQQQQQLLQQLQLQQLQLPCMDVVLQQHN 135
Db      1924  AGQVQAGVRSSTILPEQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQ--QQY- 1973

QY      136  IAHARSQVLOOSTYGLQLQELCCQHLMQIPQSQCCCAHNVVHAILHQQQKQQQQPSSQV 195
Db      1974  --HRRQQQQQQ-----QMLRQQQQQQQQQQ--QQQQQQQQQQQQQQQQQQ 2009

QY      126  SRQQPQLQVPLQGSFRPQQNPPQAQGSVQPOLQPEEIRNLA 239
Db      2010  QQQQCPHQ---QQQAAPPQPPQSQPQFQRGQLGTQQQQQTA 2049

RESULT 11
US-09-491-356C-8
; Sequence 8, Application US/09491356C
; Patent No. 6560661
; GENERAL INFORMATION:
; APPLICANT: Philibert, Robert A.
; APPLICANT: Gims, Edward I.
; APPLICANT: Delisi, Lynn
; TITLE OF INVENTION: IDENTIFICATION OF POLYMORPHISMS IN THE PCTG4 REGION OF XQ13

```

```

Query Match      14.1%; Score 202; DB 4; Length 2023;
Best Local Similarity 29.2%; Pred. No. 7.4e-10;
Matches 87; Conservative 15; Mismatches 90; Indels 106; Gaps 11;

```

```

RESULT 12
US-09-538-092-1377
; Sequence 1377, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538, 092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratSeqFormatter Version 0.9
; SEQ ID NO 1377
; LENGTH: 2124
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number Q93074
US-09-538-092-1377

```

```

Query Match 14.1% Score 202; DB 4; Length 2124;
Best Local Similarity 29.2%; Pred. No. 7,9e+10;
Matches 87; Conservative 15; Mismatches 90; Indels 106; Gaps 11
QY 7 QIQPQNPQQQOEQEVPIVQQQQQPPGQQ--QQFPQQPY----- 43
| | | | | | | | | | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | | |

```







Wed Dec 15 10:02:01 2004

us-10-089-700-3-r65.rat

Page 7

Search completed: December 14, 2004, 17:29:05  
Job time : 20 secs

---



*This Page Blank (uspto)*











```

; Sequence 101, Application US/10474955
; Publication No. US20040241161A1
; GENERAL INFORMATION:
; APPLICANT: Drifhout, Jan W.
; APPLICANT: Koning, Frits
; APPLICANT: McAdam, Stephan N.
; APPLICANT: Ludwig, Solid Magne
; TITLE OF INVENTION: METHODS AND MEANS FOR USE OF HLA-DQ RESTRICTED T-CELL RECEPTORS A
; TITLE OF INVENTION: DO BINDING PROLAMINE-DERIVED PEPTIDES
; FILE REFERENCE: 279/71244-PCT-US
; CURRENT APPLICATION NUMBER: US/10/474,955
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 101
; LENGTH: 282
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Amino acid sequence of GAMMA-1
US-10-474-955-101

```

```

Query Match          39.0%; Score 558.5; DB 17; Length 282;
Best Local Similarity 47.6%; Pred. No. 4.1e-36;
Matches 138; Conservative 33; Mismatches 72; Indels 47; Gaps 15;

```

```

QY 2 RVHPQL-----QPQNPSSQPPQSVPLVQQQPPQ--QQQF---PPQPFYPP--QPPS 51
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 10 QVMPQQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 52 --QQPFIQ--LQPPPPRLP--YPPQSPFPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 106
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 69 QPQQTYPQRPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 118
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 107 QQQQQQQQILQQILQQQLIPCMADVLLQCHN-----IAHARSQVLSQSTYLLQELCCCHLMQ 162
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 119 FPPQPPSLLIQSLQQQLNPKNPLQQCKRVSLSLSMLPSLSCQVMRQCCQQLAQ 178
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 163 IPEQSQCAIHNVVAIIILHQQKQQQPPSSQVSPQPL--QQYPLQGSFRPSQNPQAG 221
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 179 IPQQLCAAIHSIVHSIIHQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQ 229
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 222 GSVQPPQLPQFEIRNIALQTLPMQNVYIAPYCTI--APF-----GIFG 264
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 230 GIIPQPPQLPQFEIRNIALQTLPMQNVYIAPYCTI--APF-----GIFG 279
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 6
US-10-739-930-9621
; Sequence 9621, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21 (53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 9621
; LENGTH: 298
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; OTHER INFORMATION: Clone ID: TRIAB-23APR03-C125_65.P
US-10-739-930-9621

```

```

Query Match          38.3%; Score 547.5; DB 17; Length 298;
Best Local Similarity 46.9%; Pred. No. 3.2e-35;
Matches 134; Conservative 29; Mismatches 74; Indels 49; Gaps 14;
QY 4 PVPQLQPPNPSSQPPQSVPLVQQQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

DB 36 FVP--QPHQPSQPP-----QTFPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 84
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 P-QPRLPYPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 110
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 85 PQQPQPPYPPQ---QPQPPFPQTQPPQQLFPPQSQPPQPPQPPQPPQPPQPPQPPQPPQ 140
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 111 QQQILQQILQQQLIPCMADVLLQCHN-----IAHARSQVLSQSTYLLQELCCCHLMQIPEQ 166
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 141 QPPIQPSLLQQQVNPCKNPLIQCKRVSLSLSMLTPQSDQVMRQCCQQLAQIPEQ 200
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 167 SQCAIHNVVAIIILHQQKQQQPPSSQVSPQPL--QQYPLQGSFRPSQNPQAGSVQ 225
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 201 LQCAAIHIVHSIIHQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQ 251
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 226 PQQLPQFEIRNIALQTLPMQNVYIAPYCTI--APF-----GIFG 264
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 252 PQQPQLPQFEIRNIALQTLPMQNVYIAPYCTI--APF-----GIFG 297
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 7
US-10-474-955-99
; Sequence 99, Application US/10474955
; Publication No. US20040241161A1
; GENERAL INFORMATION:
; APPLICANT: Drifhout, Jan W.
; APPLICANT: Koning, Frits
; APPLICANT: McAdam, Stephan N.
; APPLICANT: Ludwig, Solid Magne
; TITLE OF INVENTION: METHODS AND MEANS FOR USE OF HLA-DQ RESTRICTED T-CELL RECEPTORS
; TITLE OF INVENTION: DO BINDING PROLAMINE-DERIVED PEPTIDES
; FILE REFERENCE: 279/71244-PCT-US
; CURRENT APPLICATION NUMBER: US/10/474,955
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 99
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of GAMMA-4
US-10-474-955-99

```

```

Query Match          37.0%; Score 529; DB 17; Length 279;
Best Local Similarity 46.5%; Pred. No. 8.6e-34;
Matches 133; Conservative 26; Mismatches 75; Indels 52; Gaps 14;

```

```

QY 4 PVPQLQPPNPSSQPPQSVPLVQQQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 18 FVP--QPHQPSQPP-----QTFPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 66
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 P-QPRLPYPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 110
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 67 PQQPQPPYPPQ---QPQPPFPQTQPPQQLFPPQSQPPQPPQPPQPPQPPQPPQPPQPPQ 122
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 111 QQQILQQILQQQLIPCMADVLLQCHN-----IAHARSQVLSQSTYLLQELCCCHLMQIPEQ 166
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 123 QPPIQPSLLQQQVNPCKNPLIQCKRVSLSLSMLTPQSDQVMRQCCQQLAQIPEQ 182
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 167 SQCAIHNVVAIIILHQQKQQQPPSSQVSPQPL--QQYPLQGSFRPSQNPQAGSVQ 225
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 183 LQCAAIHIVHSIIHQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQ 230
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 226 PQQLPQFEIRNIALQTLPMQNVYIAPYCTI--APF-----GIFG 264
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 231 PQQPQLPQFEIRNIALQTLPMQNVYIAPYCTI--APF-----GIFG 276
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 8
US-10-474-955-97
; Sequence 97, Application US/10474955
; Publication No. US20040241161A1
; GENERAL INFORMATION:

```



APPLICANT: Drifhout, Jan W.  
 APPLICANT: Konig, Frits  
 APPLICANT: McAdam, Stephan N.  
 APPLICANT: Ludwig, Soliid Magne  
 TITLE OF INVENTION: METHODS AND MEANS FOR USE OF HLA-DQ RESTRICTED T-CELL RECEPTORS A  
 FILE REFERENCE: 2799/71244-PCT-US  
 CURRENT APPLICATION NUMBER: US/10/474,955  
 CURRENT FILING DATE: 2003-10-13  
 NUMBER OF SEQ ID NOS: 137  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 97  
 LENGTH: 279  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Consensus amino acid sequence  
 US-10-474-955-97

Query Match 36.7%; Score 525; DB 17; Length 279;  
 Best Local Similarity 46.2%; Pred. No. 1.8e-33;  
 Matches 132; Conservative 27; Mismatches 75; Indels 52; Gaps 14;  
 DB 4 PVPQLQPNPSQQQPEQGVPLVQCCQFPGQQQCF--PQQPYQPQPFPSQGYLQQLPF 61  
 18 PVP--QPHQPFSSQD-----QTFPQPPQTFPHQPPQSPQPPQ--PQQFLQPPQPF 66  
 QY 62 P-QPRLPYRPPQSPFPQPPQPPQ-----PQYSPQPPPSISQQAQ-----QDQDQDQDQ 110  
 DB 67 PQQPQPPQPPQ-----QPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 122  
 QY 111 QQQILQQILQQQLIPCMQDVLLQGNH-----IAHARSQVLAQSTYQLLQELCCQHLMOIPEQ 166  
 DB 123 QPFFIQPSLQQQVNCXKFLILQCKFVLSVLSMSMTWPSQSCQVWRQSCQQLAQIIPQ 182  
 QY 167 SQCAIHNVVAIILHQQQKQQQSSQVSPQPL--QYPLGQSSFRPSQNPQAQSVQ 225  
 DB 183 LQCAIHVTHSHIIMQCEQQ-----GMHILPLVQDQVGGGTL-----VGGGGIIG 230  
 QY 226 PQQPQFEIRNALQTLPMQCNVYIAPYCTI--APF-----GIFG 264  
 DB 231 PQQPQLEAIRSLVLTPLTMQCNVYVPECSIIKAPSSVVAIGIG 276

RESULT 9  
 US-10-474-955-98  
 Sequence 98, Application US/10474955  
 Publication No. US20040241161A1  
 GENERAL INFORMATION:  
 APPLICANT: Drifhout, Jan W.  
 APPLICANT: Konig, Frits  
 APPLICANT: McAdam, Stephan N.  
 APPLICANT: Ludwig, Soliid Magne  
 TITLE OF INVENTION: METHODS AND MEANS FOR USE OF HLA-DQ RESTRICTED T-CELL RECEPTORS A  
 FILE REFERENCE: 2799/71244-PCT-US  
 CURRENT APPLICATION NUMBER: US/10/474,955  
 CURRENT FILING DATE: 2003-10-13  
 NUMBER OF SEQ ID NOS: 137  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 98  
 LENGTH: 279  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Amino acid sequence of GAMMA-2  
 US-10-474-955-98

Query Match 36.4%; Score 521; DB 17; Length 279;  
 Best Local Similarity 45.8%; Pred. No. 3.7e-33;  
 Matches 131; Conservative 27; Mismatches 76; Indels 52; Gaps 14;  
 DB 4 PVPQLQPNPSQQQPEQGVPLVQCCQFPGQQQCF--PQQPYQPQPFPSQGYLQQLPF 61

DB 18 PVP--QPHQPFSSQD-----QTFPQPPQTFPHQPPQSPQPPQ--PQQFLQPPQPF 66  
 QY 62 P-QPRLPYRPPQSPFPQPPQPPQ-----PQYSPQPPPSISQQAQ-----QDQDQDQDQ 110  
 DB 67 PQQPQPPQPPQ-----QPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 122  
 QY 111 QQQILQQILQQQLIPCMQDVLLQGNH-----IAHARSQVLAQSTYQLLQELCCQHLMOIPEQ 166  
 DB 123 QPFFIQPSLQQQVNCXKFLILQCKFVLSVLSMSMTWPSQSCQVWRQSCQQLAQIIPQ 182  
 QY 167 SQCAIHNVVAIILHQQQKQQQSSQVSPQPL--QYPLGQSSFRPSQNPQAQSVQ 225  
 DB 183 LQCAIHVTHSHIIMQCEQQ-----GMHILPLVQDQVGGGTL-----VGGGGIIG 230  
 QY 226 PQQPQFEIRNALQTLPMQCNVYIAPYCTI--APF-----GIFG 264  
 DB 231 PQQPQLEAIRSLVLTPLTMQCNVYVPECSIIKAPSSVVAIGIG 276

RESULT 10  
 US-10-474-955-100  
 Sequence 100, Application US/10474955  
 Publication No. US20040241161A1  
 GENERAL INFORMATION:  
 APPLICANT: Drifhout, Jan W.  
 APPLICANT: Konig, Frits  
 APPLICANT: McAdam, Stephan N.  
 APPLICANT: Ludwig, Soliid Magne  
 TITLE OF INVENTION: METHODS AND MEANS FOR USE OF HLA-DQ RESTRICTED T-CELL RECEPTOR  
 FILE REFERENCE: 2799/71244-PCT-US  
 CURRENT APPLICATION NUMBER: US/10/474,955  
 CURRENT FILING DATE: 2003-10-13  
 NUMBER OF SEQ ID NOS: 137  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 100  
 LENGTH: 279  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Amino acid sequence of GAMMA-3  
 US-10-474-955-100

Query Match 36.4%; Score 521; DB 17; Length 279;  
 Best Local Similarity 46.0%; Pred. No. 3.7e-33;  
 Matches 131; Conservative 29; Mismatches 75; Indels 50; Gaps 14;  
 DB 4 PVPQLQPNPSQQQPEQGVPLVQCCQFPGQQQCF--PQQPYQPQPFPSQGYLQQLPF 61  
 18 PVP--QPHQPFSSQD-----QTFPQPPQTFPHQPPQSPQPPQ--PQQFLQPPQPF 66  
 QY 62 P-QPRLPYRPPQSPFPQPPQPPQ-----PQYSPQPPPSISQQAQ-----QDQDQDQDQ 111  
 DB 67 PQQPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 123  
 QY 111 QQQILQQILQQQLIPCMQDVLLQGNH-----IAHARSQVLAQSTYQLLQELCCQHLMOIPEQ 167  
 DB 124 PFFIQPSLQQQVNCXKFLILQCKFVLSVLSMSMTWPSQSCQVWRQSCQQLAQIIPQ 183  
 QY 168 SQCAIHNVVAIILHQQQKQQQSSQVSPQPL--QYPLGQSSFRPSQNPQAQSVQ 226  
 DB 184 LQCAIHVTHSHIIMQCEQQ-----GMHILPLVQDQVGGGTL-----VGGGGIIG 231  
 QY 227 PQQPQFEIRNALQTLPMQCNVYIAPYCTI--APF-----GIFG 264  
 DB 232 PQQPQLEAIRSLVLTPLTMQCNVYVPECSIIKAPSSVVAIGIG 276

RESULT 11  
 US-10-739-930-9778  
 Sequence 9778, Application US/10739930  
 Publication No. US20040216190A1



```

GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
FILE REFERENCE: 38-21(53377)B
CURRENT APPLICATION NUMBER: US/10/739,930
CURRENT FILING DATE: 2003-12-18
NUMBER OF SEQ ID NOS: 11088
SEQ ID NO 9778
LENGTH: 304
TYPE: PRT
ORGANISM: Triticum aestivum
FEATURE:
OTHER INFORMATION: Clone ID: TRIAE-23APR03-C176_243.p
US-10-739-930-9778

```

```

Query Match
Best Local Similarity 42.1%; Score 447.5; DB 17; Length 304;
Pred. No. 2.5e-27;
Matches 123; Conservative 34; Mismatches 78; Indels 57; Gaps 16;

```

```

QY 19 QGQVPLVQQQFPQGGQQQFPQGPQYPPQPPSQQPYLQLQPP-PPRLPYPPQSGFPQ 77
DB 20 QMETSCTSGHERPQQQPLPQOSFSQPPFSQQQ--QPLPQOSFSQQQ-PFSQQ 74
QY 78 QPYPPQPPQYSPQPPISQQQAQQQQQ--QQQQQQQLLQOI-----LQQQLPQ 126
DB 75 QPILSQQPPFSQQQQPLPQOSFSQQQQLVLPQQQQQQQLVQQQIPVQPSVLQQLNFC 134
QY 127 MDVVLQGH--NIAH---ARSQVLQOSTYQLDELCCGHLWQIPESQCCAIHNVVHAI 180
DB 135 -KVFLLQCCSPVAMPORLARSCMWQSSCHVMQCCQCCQIQOIPEOSREARAIITYSII 193
QY 181 LHQQQK-----QQQPS-----SQVSFQPPQYPLQGGSPFSQNP 218
DB 194 LQEQQQGFVPPQQQPPQOSQGVSSQSQQLGQCSFPQPPQ--LQ--QFQQQQQ 248
QY 219 Q--AQGS-VOPQPPQFEIRNIALQTLPMCNVYIAPY--CTIAPFGIFGT 265
DB 249 QQVLOGTFLQPHQIAHLEAVTSIALRTLPTMCSVNVPLXATTSVDFGV-GT 299

```

```

RESULT 12
US-10-739-930-9782
Sequence 9782, Application US/10739930
Publication No. US20040216190A1
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
FILE REFERENCE: 38-21(53377)B
CURRENT APPLICATION NUMBER: US/10/739,930
CURRENT FILING DATE: 2003-12-18
NUMBER OF SEQ ID NOS: 11088
SEQ ID NO 9782
LENGTH: 307
TYPE: PRT
ORGANISM: Triticum aestivum
FEATURE:
OTHER INFORMATION: Clone ID: TRIAE-23APR03-C176_269.p
US-10-739-930-9782

```

```

Query Match
Best Local Similarity 42.8%; Score 440.5; DB 17; Length 307;
Pred. No. 9.1e-27;
Matches 125; Conservative 37; Mismatches 75; Indels 55; Gaps 16;

```

```

QY 115 -LQQLLQOQLPMDVVLQGH--NIAH---ARSQVLQOSTYQLDELCCGHLWQIPES 167
DB 138 VQPSILQQLNFC-KVFLLQCCSPVAMPORLARSCMWQSSCHVMQCCQCCQQLPQPS 196
QY 168 QCOAIHNVVHAIILHQQKQQQOSQVSFQOPLQOYPLQGGSPFSQNPQ-----219
DB 197 REARAIITYSIIL--QEQQVQSGISQQQQPPQ--LGCVSQPPQOSQQLGQCPQ 251
QY 220 ---AQGS-VOPQPPQFEIRNIALQTLPMCNVYIAPY--CTIAPFGIFGT 265
DB 252 QQLAQGTFLQPHQIAHLEAVTSIALRTLPTMCSVNVPLXATTSVDFGV-GT 302

```

```

RESULT 13
US-10-739-930-9769
Sequence 9769, Application US/10739930
Publication No. US20040216190A1
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
FILE REFERENCE: 38-21(53377)B
CURRENT APPLICATION NUMBER: US/10/739,930
CURRENT FILING DATE: 2003-12-18
NUMBER OF SEQ ID NOS: 11088
SEQ ID NO 9769
LENGTH: 244
TYPE: PRT
ORGANISM: Triticum aestivum
FEATURE:
OTHER INFORMATION: Clone ID: TRIAE-23APR03-C176_124.p
US-10-739-930-9769

```

```

Query Match
Best Local Similarity 40.3%; Score 384; DB 17; Length 244;
Pred. No. 2e-22;
Matches 108; Conservative 34; Mismatches 56; Indels 70; Gaps 14;

```

```

QY 39 PQQPYP-QPP--FPGQPYLQLQPPQPLPYPPQPSF-----PQ 77
DB 1 PQQPFLQPPQSFQMSQPPQ--QFQPSFPQVQVQIIPATPTTIPAGKPTS 55
QY 78 QPYPPQPPQYSPQPPISQQQAQQQQQ--QQQQQQQLLQOIPLPMDVVLQGH--N 135
DB 56 APFPQ-----QQQHQQLAQQOI PVQPSILQQLNFC-KVFLLQCCSP 97
QY 136 IAH---ARSQVLQOSTYQLDELCCGHLWQIPESQCCAIHNVVHAIILHQQKQQQ 191
DB 98 VAMPORLARSCMWQSSCHVMQCCQCCQQLPQIPQSRQAIITYSIIL--QEQQVQ 155
QY 192 SSQVSFQOPLQOYPLQGGSPFSQNPQ-----AQGS-VOPQPPQFEIRNIA 239
DB 156 SIQSQQQPPQ--LGCVSQPPQOSQQLGQCPQOQLAQGTFLQPHQIAHLEAVTSIA 212
QY 240 LQTLPMCNVYIAPY--CTIAPFGIFGT 265
DB 213 LRILPTMCSVNVPLXATTSVDFGV-GT 239

```

```

RESULT 14
US-10-425-115-200100
Sequence 200100, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Cao, Yongwei
APPLICANT: Zhou, Yihua
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326

```















Best Local Similarity 84.7%; Pred. No. 1.9e-77;  
Matches 243; Conservative 9; Mismatches 14; Indels 21; Gaps 3;

```

QY 1 VRVPVPOLOPQONPSQOQPOEQVPLVQOQOQPPGQOQOQPPPOQPYPOQPPPSQOQPYLOLP 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 21 VRVPVPOLOPQONPSQOQPOEQVPLVQOQOQPPGQOQOQPPPOQPYPOQPPPSQOQPYLOLP 80
QY 61 F-----PQRLPYPOQSPFPPOQPYPOQPOQYQSQOQPISSQOQAQOQOQOQ 105
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 81 FQOPQPYPOQLOPQPYPOQLOPQPYPOQPYPOQPYPOQPYPOQPYPOQPYPOQPYPOQ 140
QY 106 --QOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ 163
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 141 QKQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ 200
QY 164 PEOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ 219
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 201 PEOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ 260
QY 220 AOGSVOPQOLPQFEIRNLALQTLPMCNVYIAPYCTTIAPFGIFGTN 266
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 261 AOGSVOPQOLPQFEIRNLALQTLPMCNVYIAPYCTTIAPFGIFGTN 307

```

## RESULT 6

507361  
alpha/beta-gliadin precursor (clone pm215) - wheat  
C/Species: Triticum aestivum (common wheat)  
C/Date: 08-Jun-1994 #sequence\_revision 01-Dec-1995 #text\_change 09-Jul-2004  
C/Accession: S07361  
R:Summer-Smith, M.; Rafalski, J.A.; Sugiyama, T.; Stoll, M.; Seell, D.  
Nucleic Acids Res. 13, 3905-3916, 1985  
A/Title: Conservation and variability of wheat alpha/beta-gliadin genes.  
A/Reference number: S07361; PMID:85242077; PMID:3839304  
A/Accession: S07361  
A/Status: preliminary  
A/Molecule type: DNA  
A:Residues: 1-296 <SUM>  
A:Cross-references: UNIPROT:P04726; EMBL:X02539; NID:g21756; PIDD:CAA26383.1; PID:g21757  
C/Superfamily: gliadin  
C/Keywords: seed; storage protein

Query Match 88.3%; Score 1263; DB 2; Length 296;  
Best Local Similarity 88.1%; Pred. No. 3.6e-77;  
Matches 244; Conservative 5; Mismatches 16; Indels 12; Gaps 4;

```

QY 1 VRVPVPOLOPQONPSQOQPOEQVPLVQOQOQPPGQOQOQPPPOQPYPOQPPPSQOQPYLOLP 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 21 VRVPVPOLOPQONPSQOQPOEQVPLVQOQOQPPGQOQOQPPPOQPYPOQPPPSQOQPYLOLP 80
QY 61 FPO-----PRLPYPOQSPFPPOQPYPOQPOQYQSQOQPISSQOQAQOQOQOQOQOQ 111
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 81 FPOQPPPPPOQPYPOQPPPSQOQPYPOQPYPOQPYPOQPYPOQPYPOQPYPOQPYPOQ 140
QY 112 QOQILOQLOQOQILPCMDVVLQOHNIAHARSQVLOQSTYQLLQELCCQHLMOIPEQSCQCA 171
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 141 QOQILOQLOQOQILPCMDVVLQOHNIAHARSQVLOQSTYQLLQELCCQHLMOIPEQSCQCA 200
QY 172 INHVVAHILHQQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ 231
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 201 INHVVAHILHQQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ 259
QY 232 FEEIRNLALQTLPMCNVYIAPYCTTIAPFGIFGTN 266
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 260 FEEIRNLALQTLPMCNVYIAPYCTTIAPFGIFGTN 296

```

## RESULT 7

507924  
alpha/beta-gliadin precursor - wheat  
C/Species: Triticum aestivum (common wheat)  
C/Date: 08-Jun-1994 #sequence\_revision 01-Dec-1995 #text\_change 09-Jul-2004  
C/Accession: S07924; C61218

R:Summer-Smith, M.; Rafalski, J.A.; Sugiyama, T.; Stoll, M.; Seell, D.  
Nucleic Acids Res. 13, 3905-3916, 1985  
A/Title: Conservation and variability of wheat alpha/beta-gliadin genes.  
A/Reference number: S07361; PMID:85242077; PMID:3839304  
A/Accession: S07361  
A/Status: preliminary; translation not shown  
A/Molecule type: DNA  
A:Residues: 1-313 <SUM>  
A:Cross-references: UNIPROT:Q41546; EMBL:X02540; NID:g21764; PIDD:CAA26385.1; PID:g2176  
R:Shewry, P.R.; Sabelli, P.A.; Parmar, S.; Lalitandra, D.  
Biochem. Genet. 29, 207-211, 1991  
A/Title: alpha-type proteins are encoded by genes on chromosomes 4A and 6A of Haynall  
A/Reference number: A61218; PMID:91315394; PMID:1859356  
A/Accession: C61218  
A/Status: preliminary  
A/Molecule type: protein  
A:Residues: 18-27 <SHE>  
C/Superfamily: gliadin  
C/Keywords: seed; storage protein

Query Match 88.1%; Score 1261; DB 2; Length 313;  
Best Local Similarity 82.8%; Pred. No. 5.2e-77;  
Matches 245; Conservative 9; Mismatches 12; Indels 30; Gaps 4;

```

QY 1 VRVPVPOLOPQONPSQOQPOEQVPLVQOQOQPPGQOQOQPPPOQPYPOQPPPSQOQPYLOLP 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 18 VRVPVPOLOPQONPSQOQPOEQVPLVQOQOQPPGQOQOQPPPOQPYPOQPPPSQOQPYLOLP 77
QY 61 FPO-----PRLPYPOQSPFPPOQPYPOQPOQYQSQOQPISSQOQAQOQOQOQOQOQOQ 112
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 78 FPOQPPPPPOQPYPOQPPPSQOQPYPOQPYPOQPYPOQPYPOQPYPOQPYPOQPYPOQ 137
QY 113 QILOQILOQLOQILPCMDVVLQOHNIAHARSQVLOQSTYQLLQELCCQHLMOIPEQSCQCA 172
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 138 QILOQILOQLOQILPCMDVVLQOHNIAHARSQVLOQSTYQLLQELCCQHLMOIPEQSCQCA 197
QY 173 HNVVAHILH-----QOQKQOQOQSSQVSSFOQPLQOYPLGQGSFR 212
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 198 HNVVAHILHMQEQEQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQOQ 257
QY 213 PSQONPQOQSVOPQOLPQFEIRNLALQTLPMCNVYIAPYCTTIAPFGIFGTN 266
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 258 PSQONPQOQSVOPQOLPQFEIRNLALQTLPMCNVYIAPYCTTIAPFGIFGTN 313

```

## RESULT 8

A27319  
gliadin - wheat  
C/Species: Triticum aestivum (common wheat)  
C/Date: 04-Mar-1988 #sequence\_revision 04-Mar-1988 #text\_change 03-Feb-1994  
C/Accession: A27319  
R:Reeves, C.D.; Okita, T.W.  
Gene 52, 257-266, 1987  
A/Title: Analyses of alpha/beta-type gliadin genes from diploid and hexaploid wheats.  
A/Reference number: A27319; PMID:87273398; PMID:3038688  
A/Accession: A27319  
A/Status: preliminary  
A/Molecule type: DNA  
A:Residues: 1-296 <REB>  
C/Superfamily: gliadin

Query Match 87.8%; Score 1257; DB 2; Length 296;  
Best Local Similarity 87.7%; Pred. No. 9.1e-77;  
Matches 243; Conservative 4; Mismatches 18; Indels 12; Gaps 4;

```

QY 1 VRVPVPOLOPQONPSQOQPOEQVPLVQOQOQPPGQOQOQPPPOQPYPOQPPPSQOQPYLOLP 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 21 VRVPVPOLOPQONPSQOQPOEQVPLVQOQOQPPGQOQOQPPPOQPYPOQPPPSQOQPYLOLP 80
QY 61 FPO-----PRLPYPOQSPFPPOQPYPOQPOQYQSQOQPISSQOQAQOQOQOQOQOQOQ 111
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 81 FPOQPPPPPOQPYPOQPPPSQOQPYPOQPYPOQPYPOQPYPOQPYPOQPYPOQPYPOQ 140

```















QY	VRVVEPDLQGNQNSQQQPOBEQVYLWQOQQPFCQQQQQFPQQQVYPOQCFPFQQQPYLQIQP	60
Db	VRVVEPDLQGNQNSQQQPOBEQVYLWQOQQPFCQQQQPFPQQQVYPOQCFPFQQQPYLQIQP	61
QY	61 PPOPRFLYPOQSFPPOQYPIPOBPQYPOQPISQQAQQAQQQQQQQQQQQQLIQQLQ	120
Db	62 FPOQFLYPSQPQFFRFQOPYPQBPQPSQPPISQQAQQQQQQQQQQQQQLIQQLQ	121
QY	121 QQLIPCDVVLQOQHNIHAHRSQVLAQSTYQLLQELCCOHLWQIPQSQCOAHNVVAHII	180
Db	122 QQLIPCDVVLQOQHNIHAHRSQVLAQSTYQLLQELCCOHLWQIPQSQCOAHNVVAHII	181
QY	181 LHQQQKQOQOQPSQVQSFQOPLQQYPIGQSSFPSSQONPQAQGSVAPQQLPQFEERINLAL	240
Db	182 LHQQQKQOQOQPSQVQSFQOPLQQYPIGQSSFPSSQONPQAQGSVAPQQLPQFEERINLAL	241
QY	241 QTLPAQNCVYIAEYCTIAPFGIFGTN	266
Db	242 QTLPAQNCVYIIPYCTIAPFGIFGTN	267



## RESULT 2

GENO WHEAT STANDARD; PRT; 286 AA.  
 ID: GDAO\_WHEAT P02863;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Alpha/beta-gliadin precursor (Prolamins).  
 OS Triticum aestivum (Wheat).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
 OC Triticeae; Triticum.  
 NC NCBI\_Taxid=4565;  
 RN 11  
 RS SEQUENCE FROM N.A.  
 RC STRAIN=cv. Newton;  
 RA MEDLINE=8426134; PubMed=6204862;  
 RA Rafalski J.A., Scheets K., Metzler M., Peterson D.M., Hedgcock C.,  
 RA Soll D.G.;  
 RA "Developmentally regulated plant genes: the nucleotide sequence of a  
 RA wheat gliadin genomic clone."  
 RA EMBO J. 3:1409-1415(1984).  
 RN 12  
 RS SEQUENCE FROM N.A. (CLONE PM8233).  
 RA MEDLINE=85242077; PubMed=3839304;  
 RA Summer-Smith M., Rafalski J.A., Sugiyama T., Scoll M., Soell D.;  
 RA "Conservation and variability of wheat alpha/beta-gliadin genes."  
 RA Nucleic Acids Res. 13:3905-3916(1985).  
 RN 13  
 RS SEQUENCE FROM N.A.  
 RA MEDLINE=85062803; PubMed=6095191;  
 RA Anderson O.D., Litts J.C., Gautier M.F., Greene F.C.;  
 RA "Nucleic acid sequence and chromosome assignment of a wheat storage  
 RA protein gene."  
 RA Nucleic Acids Res. 12:8129-8144(1984).  
 CC -1- FUNCTION: Gliadin is the major seed storage protein in wheat.  
 CC -1- PRT: Substrate of transglutaminase (By similarity).  
 CC -1- ALLERGEN: Causes an allergic reaction in human. Is the cause of  
 CC the celiac disease, also known as celiac sprue or gluten-sensitive  
 CC enteropathy (By similarity).  
 CC -1- MISCELLANEOUS: The alpha/beta-gliadins can be divided into 5  
 CC homology classes. Sequence divergence between the classes is due  
 CC to single base substitutions and to duplications or deletions  
 CC within or near direct repeats. There are more than a 100 copies of  
 CC the gene for alpha/beta-gliadin per haploid genome.  
 CC  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL: X00627; CAA25261.1; -  
 CC EMBL: X03076; AAA34280.1; -  
 CC EMBL: X02539; CAA26384.1; -  
 CC EMBL: X01130; CAA25583.1; -  
 CC PIR: A03354; EEMTH.  
 CC InterPro: IPR003612; AAI.  
 CC InterPro: IPR001376; Gliadin.  
 CC Pfam: PF00234; TTRP\_alpha\_amy1.1.  
 CC PRINTS: PR00208; GLIADGLUTEN.  
 CC PRINTS: PR00209; GLIADIN.  
 CC SMART: SM00499; AAI; 1.  
 CC KX Allergen: Multigene family; Repeat; Seed storage protein; Signal.  
 CC CHAIN 1 20  
 CC SIGNAL 1 20  
 CC VARIANT 37 37 L -> Q (in clone PM8233 and in Ref. 3).  
 CC VARIANT 93 93 P -> Q (in clone PM8233).

FT VARIANT 193 194 HN -> LR (in Ref. 3).  
 SQ SEQUENCE 286 AA; 32949 MW; E5ECFABBE29E10C6 CRC64;

Query Match 94.8%; Score 1356; DB 1; Length 286;  
 Best Local Similarity 95.5%; Pred. No. 6.9e-75;  
 Matches 254; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

QY 1 VRVPVQLQPNPSSQQLPQEQVPLVQGGQFPQGGQFPQGPYPQGPFPSSQGYLQQLP 60  
 DB 21 VRFPVQLQPNPSSQQLPQEQVPLVQGGQFPQGGQFPQGPYPQGPFPSSQGYLQQLP 80  
 QY 61 FPGPRLPYQPQSPFPQGPYPQGPYPQGPYPQGPYPQGPYPQGPYPQGPYPQGPYPQGP 120  
 DB 81 FPGPRLPYQPQSPFPQGPYPQGPYPQGPYPQGPYPQGPYPQGPYPQGPYPQGPYPQGP 140  
 QY 121 QQLIPGMDVYLQOHNIHARSOVLQOSTYQLQELCCOHLWQIPQSGCCQAIHNVVHAII 180  
 DB 141 QQLIPGMDVYLQOHNIHARSOVLQOSTYQLQELCCOHLWQIPQSGCCQAIHNVVHAII 200  
 QY 181 LHQGQKQKQKQSSQVSPFQGLQGYPLGGQSPFPSSQGNPQAGSVQPPQLPQFEETRNAL 240  
 DB 201 LHQGQKQKQKQSSQVSPFQGLQGYPLGGQSPFPSSQGNPQAGSVQPPQLPQFEETRNAL 260  
 QY 241 QTLPMQNVYIAPYCTIAPFGIGTN 266  
 DB 261 QTLPMQNVYIAPYCTIAPFGIGTN 286

RESULT 3  
 ID AAA96525 PRELIMINARY; PRT; 286 AA.  
 AC AAA96525;  
 DT 02-MAR-2004 (TRENBLrel. 27, Created)  
 DT 02-MAR-2004 (TRENBLrel. 27, Last sequence update)  
 DT 02-MAR-2004 (TRENBLrel. 27, Last annotation update)  
 DE Alpha-gliadin storage protein.  
 OS Triticum aestivum (Wheat).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Triticeum.  
 NC NCBI\_Taxid=4565;  
 RN 11  
 RS SEQUENCE FROM N.A.  
 RC STRAIN=Cheyenne;  
 RA Anderson O.D.;  
 RA Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U51307; AAA96525.1; -  
 SQ SEQUENCE 286 AA; 32949 MW; E5ECFABBE29E10C6 CRC64;

Query Match 94.8%; Score 1356; DB 2; Length 286;  
 Best Local Similarity 95.5%; Pred. No. 6.9e-75;  
 Matches 254; Conservative 1; Mismatches 11; Indels 0; Gaps 0;  
 QY 1 VRVPVQLQPNPSSQQLPQEQVPLVQGGQFPQGGQFPQGPYPQGPFPSSQGYLQQLP 60  
 DB 21 VRFPVQLQPNPSSQQLPQEQVPLVQGGQFPQGGQFPQGPYPQGPFPSSQGYLQQLP 80  
 QY 61 FPGPRLPYQPQSPFPQGPYPQGPYPQGPYPQGPYPQGPYPQGPYPQGPYPQGPYPQGP 120  
 DB 81 FPGPRLPYQPQSPFPQGPYPQGPYPQGPYPQGPYPQGPYPQGPYPQGPYPQGPYPQGP 140  
 QY 121 QQLIPGMDVYLQOHNIHARSOVLQOSTYQLQELCCOHLWQIPQSGCCQAIHNVVHAII 180  
 DB 141 QQLIPGMDVYLQOHNIHARSOVLQOSTYQLQELCCOHLWQIPQSGCCQAIHNVVHAII 200  
 QY 181 LHQGQKQKQKQSSQVSPFQGLQGYPLGGQSPFPSSQGNPQAGSVQPPQLPQFEETRNAL 240  
 DB 201 LHQGQKQKQKQSSQVSPFQGLQGYPLGGQSPFPSSQGNPQAGSVQPPQLPQFEETRNAL 260  
 QY 241 QTLPMQNVYIAPYCTIAPFGIGTN 266  
 DB 261 QTLPMQNVYIAPYCTIAPFGIGTN 286



OX	NCBI_TaxID=4565;												
RN	[1]												
RP	SEQUENCE FROM N.A.												
RC	STRAIN=Mjoeiner; TISSUE=Endosperm;												
RA	Arentz-Hansen E.H., McAdam S.N., Molberg O., Kristiansen C.,												
RL	Sollid L.M.;												
DR	Submitted (MAR-1999) to the EMBL/genbank/DBJ databases.												
DR	EMBL: AJ133603; CAB76955.1;												
DR	GO: GO:0045735; F:nutrient reservoir activity; IEA.												
DR	InterPro: IPR003612; AAT												
DR	InterPro: IPR001376; Gliadin.												
DR	Pfam: PF00234; TRYD_alpha.amyl.1.												
DR	PRINTS: PR00208; GLIADGLUTEN.												
DR	PRINTS: PR00209; GLIADIN.												
DR	SMART: SM00499; AAI; 1.												
FT	CHAIN 1 274 alpha-gliadin.												
EQ	SEQUENCE 274 AA; 31980 MW; 976919397534ABD CRC64;												
Query Match													
Best Local Similarity 94.4%; Score 1350.5; DB 2; Length 274;													
Matches 255; Conservative 1; Mismatches 10; Indels 5; Gaps 1													
DQ	1	VRYVYQLDPQPNSQQQPEQVYLVOOQOQFQGQQQAFPPQCPYPQCPFPSPGQPYLOLP	60										
DQ	2	VRVTVAQLDPQNPSQQQPEQVYLVOOQOQFQGQQQFPFPQCPYPQCPFPSPGQPYLOLP	61										
DQ	61	FPGPRLHPYQPQSFPFQCPYPQCPQYSQPQPS-----QQQAQQQQQQQQQQQQQQL	115										
DQ	62	FPGPQLPYSPQPFRRQCPYPQCPQYSQPQGPISQQQQQQQQQQQQQQQQQQQQQL	121										
DQ	116	QQILQQQLIPCMDVVLQCHNIHANSQVLQGSTTQLCELCCQHLMQIPBSCCAIHNV	175										
DQ	122	QQILQQQLIPCMDVVLLQCHNIHGRSQVLQGSTVQLLELCCHLMQIPESCOAHNV	181										
DQ	176	VHAIIHHQKKKQQSSQVSQCPLOOQYPIQGSGFRSQONPONAQSVAPQQLPQEEL	233										
DQ	182	VHAIIHHQKKKQQSSQVSQCPLOOQYPIQGSGFRSQONPONAQSVAPQQLPQEEL	241										
DQ	236	RNLALQTLPAMCNVYIAPYCTIAPFGIFGTN	266										
DQ	242	RNLALQTLPAMCNVYIAPYCTIAPFGIFGTN	272										
RESULT 6													
ID	Q9M4M2	PRELIMINARY;	PRT;	276	AA.								
AC	G3M4M2;												
DT	01-OCT-2000 (TrEMBLrel. 15, Created)												
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)												
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)												
DE	Alpha-gliadin.												
OS	Triticum aestivum (wheat).												
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;												
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;												
OC	Triticeae; Triticum.												
CX	NCBI_TaxID=4565;												
RN	[1]												
RP	SEQUENCE FROM N.A.												
RC	STRAIN=Mjoeiner; TISSUE=Endosperm;												
RA	Arentz-Hansen E.H., McAdam S.N., Molberg O., Kristiansen C.,												
RL	Sollid L.M.;												
DR	Submitted (MAR-1999) to the EMBL/genbank/DBJ databases.												
DR	EMBL: AJ133603; CAB76955.1;												
DR	GO: GO:0045735; F:nutrient reservoir activity; IEA.												
DR	InterPro: IPR003612; AAT												
DR	InterPro: IPR001376; Gliadin.												
DR	Pfam: PF00234; TRYD_alpha.amyl.1.												
DR	PRINTS: PR00208; GLIADGLUTEN.												
DR	PRINTS: PR00209; GLIADIN.												



SQ	SEQUENCE	276 AA;	32213 MW;	6A2CDDCAE70BA100A CRC64;
	Query Match	94.4%;	Score 1350.5;	DB 2; Length 276;
	Best Local Similarity	93.4%;	Pred. No. 1.4e-74;	
	Matches 255; Conservative	2;	Mismatches 9;	Indels 7; Gaps 1;
OY	1 VRVVPOLPQPONPSQQQDEQVPLVQQQGFPCGQQQFPPGQGPYPGQPFPSQQPYLQLQP	60		
DB	2 VRVVPQLQPQNPSQQQDEQVPLVQQQGLGQQQFPFGQGPYPGQPFPSQQPYLQLQP	61		
OY	61 PPQPLPLPFGQSPFPQQPYPGQPFPSQQQDPTSSQQQAAGQQQQQQQQQQ-----QQ	113		
DB	62 PPQPLPLSQQQPFPPQQPYPGQPFPSQQQDPTSSQQQAAGQQQQQQQQQQ-----QQ	121		
OY	114 ILQQLILQQLIPCMNVVLQGNHIAHRSQVLQQSTYLQELCCGHLMQIPEQSQQAIH	173		
DB	122 ILQQLILQQLIPCMNVVLQGNHIAHRSQVLQQSTYLQELCCGHLMQIPEQSQQAIH	181		
OY	174 NVTEALITHHQKQKQQQSSSQVSFPQPILOQVYLGGSGRRSPQQNDAQGSVQPOQLPQE	233		
DB	182 NVVAHIILHHQDKQQQSSSQVSFPQPILOQVPLGGSGRRSPQQNPDAQGSVQPOQLPQE	241		
OY	234 EIRNALQTLPMACNVVIAPYCTTAPFGIFGTN	266		
DB	242 EIRNALQTLPMACNVVIAPYCTTAPFGIFGTN	274		
<hr/>				
	RESULT 7			
TO	G41509	PRELIMINARY;	FRT;	287 AA.
AC	G41509			
DT	01-NOV-1996 (TREMBLrel. 01, Created)			
DI	01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
DL	01-MAR-2004 (TREMBLrel. 26, Last annotation update)			
DR	Alpha-Gladin.			
OS	Triticum aestivum (Wheat).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OG	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;			
OX	Triticeae; Triticum.			
XX	NCBI_TaxID=4565;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Cheyenne;			
RA	Blechl A.E., Anderson O.D.;			
RL	Submitted (Apr-1994) to the EMBL/GenBank/DDBJ databases.			
DR	EMBL; U08287; AAA17741.1; -.			
OR	PIR; S1333; S1333.			
DR	GO; GO:0045735; Pnutrient reservoir activity; IEA.			
DR	InterPro; IPR003612; AAI.			
DR	InterPro; IPR001376; Gliadin.			
DR	InterPro; IPR001954; Glia_Glutenin.			
DR	pfam; PF00234; TRY_P_alpha_amy1; 1.			
DR	PRINTS; PR00208; GLIADINUTEN.			
DR	PRINTS; PR00209; GLIADIN.			
DR	SMART; SM00499; AAI; 1.			
SQ	SEQUENCE	287 AA;	BF980C21CFB84873 CRC64;	
<hr/>				
	Query Match	94.0%;	Score 1345.5;	DB 2; Length 287;
	Best Local Similarity	94.8%;	Pred. No. 3e-74;	
	Matches 253; Conservative	3;	Mismatches 10;	Indels 1; Gaps 1
OY	1 VVVVPVLPOLPONPSQQQDEQVPLVQQQGFPPGQQQQFPFGQGPYPGQPFPSQQPYLQLQP	60		
DB	21 VVVVPVLPOLPONPSQQQDEQVPLVQQQGLGQQQFPFGQGPYPGQPFPSQQPYLQLQP	60		
OY	61 PPQPLPLPFGQSPFPQQPYPGQPFPSQQQDPTSSQQQAAGQQQQQQQQQQ-----QQ	119		
DB	81 PPQPLPLPFGQSPFPQQPYPGQPFPSQQQDPTSSQQQAAGQQQQQQQQQQ-----QQ	140		
OY	120 QOQLITCMNVVLQGNHIAHRSQVLQQSTYLQELCCGHLMQIPEQSQQAIHVVAHI	179		
DB	141 QOQLITCMNVVLQGNHIAHRSQVLQQSTYLQELCCGHLMQIPEQSQQAIHVVAHI	200		

QY	180	ILHQQQKQQQQSSQVSVFQQPLQQYPLQGQSFRRSQNPPAQGQSVQQLPQFEEIINLA	233
Db	201	ILHQQQKQQQQSSQVSVFQQPLQQYPLQGQSFRRSQNPPAQGQSVQQLPQFEEIINLA	260
QY	240	LQTLPMQCNVYIAPYCTIAPPGIFGTN	266
Db	261	LQTLPMQCNVYIAPYCTIAPPGIFGTN	287
RESULT 8			
QY	Q9M4L8	PRELIMINARY;	PRT; 277 AA.
AC	Q9M4L8		
DT	01-OCT-2000	(TREMBLrel. 15, Created)	
DT	01-OCT-2000	(TREMBLrel. 15, Last sequence update)	
DT	01-MAR-2004	(TREMBLrel. 26, Last annotation update)	
DE	Alpha-glutadin.		
OS	Triticum aestivum (wheat).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
CC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;		
CC	Triticeae; Triticum.		
CX	NCBI_Taxid=4565;		
RL	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Mjcelner; TISSUE=Endosperm;		
RA	Arantz-Hansen E.H., Wcadam S.N., Molberg O., Kristiansen C.,		
RA	Sollid L.M.;		
RL	Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AJ133610; CAB76962.1; -		
DR	GO: GO:0045735; F:nutrient reservoir activity; IEA.		
DR	InterPro; IPR003612; AAI.		
DR	InterPro; IPR001376; Glutadin.		
DR	InterPro; IPR001954; Glia_glutenin.		
DR	Pfam; PF00234; TRYP_alpha_amy; 1.		
DR	PRINTS; PR00208; GLIADGLUTEN.		
DR	PRINTS; PR00209; GLIADIN.		
DR	SMART; SMO0499; AAI; 1.		
FT	CHAIN 1	alpha-glutadin.	
FT	SEQUENCE 277 AA; 32371 MW; 73DB85D815E5329D CRC64;		
QY	Query Match	93.9%; Score 1344; DB 2; Length 277;	
	Best Local Similarity	92.7%; Pred. No. 3,6e-74;	
	Matches 254; Conservative 1; Mismatches 11; Indels 9; Gaps 1		
QY	1	VRVPRPQIQPQNSQCPQCBQVPLVQVQVQFQGGQQQFPQQPQYPQPPQFPSSQAPYLQDP	60
Db	2	VRVPRPQIQPQNSQCPQCBQVPLVQVQVQFQGGQQQFPQQPQYPQPPQFPSSQAPYLQDP	61
QY	61	FPQPRPLPQCPQSPFPQQPQYPQPPQFQVQSQPQPIQ-----QQQAQQQQQQQQQQQQ	112
Db	62	FPQPRPLPQCPQSPFPQQPQYPQPPQFQVQSQPQPIQ-----QQQAQQQQQQQQQQQQ	121
QY	113	QILQQILQQQILPCMDVVLQGHNINHAHQVYQCTYQLLQELCCGHLMQVPEQSQQCAI	172
Db	122	QILQQILQQQILPCMDVVLQGHNINHAHQVYQCTYQLLQELCCGHLMQVPEQSQQCAI	181
QY	173	HNHVAHIIHQQQKQQQQSSQVSVFQQPLQQYPLQGQSFRRSQNPPAQGQSVQQLPQF	232
Db	182	HNHVAHIIHQQQKQQQQSSQVSVFQQPLQQYPLQGQSFRRSQNPPAQGQSVQQLPQF	241
QY	233	EEIRNIALQTLPMQCNVYIAPYCTIAPPGIFGTN	266
Db	242	EEIRNIALQTLPMQCNVYIAPYCTIAPPGIFGTN	275
RESULT 9			
QY	Q9M4M0	PRELIMINARY;	PRT; 276 AA.
AC	Q9M4M0		
DT	01-OCT-2000	(TREMBLrel. 15, Created)	
DT	01-OCT-2000	(TREMBLrel. 15, Last sequence update)	
DT	01-MAR-2004	(TREMBLrel. 26, Last annotation update)	
DE	Alpha-glutadin.		



[illegible]

Query Match	93.5%	Score 1337.5	DB 2	Length 289
Best Local Similarity	94.1%	Pred. No. 9.2e-74		
Matches 253	Conservative 2	Mismatches 11	Indels 3	Gaps 1
SO	SEQUENCE	289 AA	33349 MW	5F577C9CD63874FA CRC64
DR	PRINTS	PR00208	GLIADGLUTEN	
DR	PRINTS	PR00209	GLIADIN	
DR	SMART	SM00499	AAI: 1	
DR	SEQUENCE	289 AA	33349 MW	5F577C9CD63874FA CRC64
QY	1	VRVAVPOLQPNPNCQQQPOQEVPLVVOQQQFFGQQQQQFPFQQQPYQPQPFPSQQQPYLQIQ	60	
DB	21	VRVAVPOLQPNPNCQQQPOQEVPLVVOQQQFFGQQQQQFPFQQQPYQPQPFPSQQQPYLQIQ	80	
QY	61	FPQRLPFPQPSFPQCPYQPOQPOYSQPCQPIIS---QQQACQQQQQQQQQQQQQQQQQLIQ	117	
DB	81	FPQRLPFPQPSFPQCPYQPOQPOYSQPCQPIISQQQQQQQQQQQQQQQQQQQLIQ	140	
QY	118	ILQQQLIPCDVYVQGHNIHARQQVLAQSTYQLLOELCCGHLWCIPESQSCQAIHNVH	177	
DB	141	ILQQQLIPCDVYVQGHNIHARQQVLAQSTYQLLOELCCGHLWCIPESQSCQAIHNVH	200	
QY	178	AIILHQOQKQQQSSQVSPQCPILQXVPLGQGSFRPSSQNPQAOQSVQPOOLPQEEIRN	237	
DB	201	AIILHQOQKQQQSSQVSPQCPILQXVPLGQGSFRPSSQNPQAOQSVQPOOLPQEEIRN	260	
QY	238	LALQTLPMCMNVYIAPYCTIAPFGIFGTN	266	
DB	261	LALQTLPMCMNVYIAPYCTIAPFGIFGTN	289	
RESULT 11				
Q9M4L9	PRELIMINARY	PRT	270 AA	
ID	Q9M4L9			
AC	Q9M4L9			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, last sequence update)			
DT	01-MAR-2004 (TREMBLrel. 26, last annotation update)			
DT	Alpha-gladiin.			
OS	Triticum aestivum (wheat).			
OC	Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta,			
CC	Spermatophyta, Magnoliophyta, Liliopsida, Poales, Poaceae, Pooidae,			
CC	Triticeae, Triticum.			
OX	NCBI_TaxID=4565;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	STRAIN=Mjoelner; TISSUE=Endosperm;			
RA	Arentz-Hansen E.H., McAdam S.N., Molberg O., Kristiansen C.,			
RA	Solid L.M.;			
RL	Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.			
DR	EMBL, AJ133609; CAB76961.1; --			
DR	GO; GO:0045735; F:nutrient reservoir activity; IEA.			
DR	InterPro; IPR003612; AAI.			
DR	InterPro; IPR001376; Gliadin.			
DR	InterPro; IPR001954; GliA glutenin.			
DR	Pfam; PF00234; T1yp_alpha_amy1; 1.			
DR	PRINTS; PR00208; GLIADGLUTEN.			
DR	PRINTS; PR00209; GLIADIN.			
DR	SMART; SM00499; AAI: 1.			
FT	CHAIN 1 270 alpha-gladiin.			
SQ	SEQUENCE 270 AA; 31491 MW; IDB4B626EFPADF5 CRC64;			
Query Match	93.0%	Score 1330.5	DB 2	Length 270
Best Local Similarity	94.0%	Pred. No. 2.3e-73		
Matches 251	Conservative 4	Mismatches 11	Indels 1	Gaps 1
QY	1	VRVAVPOLQPNPNCQQQPOQEVPLVVOQQQFFGQQQQQFPFQQQPYQPQPFPSQQQPYLQIQ	60	
DB	2	VRVAVPOLQPNPNCQQQPOQEVPLVVOQQQFFGQQQQQFPFQQQPYQPQPFPSQQQPYLQIQ	61	
QY	61	FPQRLPFPQPSFPQCPYQPOQPOYSQPCQPIISQQQACQQQQQQQQQQQQQQQQQLIQ	119	
DB	62	FPQRLPFPQPSFPQCPYQPOQPOYSQPCQPIISQQQACQQQQQQQQQQQQQQQQQLIQ	121	



[illegible]

PT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Alpha-gliadin (Fragment).  
 OS Triticum aestivum (Wheat).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
 OC Triticaceae; Triticum.  
 OC NCBI\_TaxID=4565;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RP MEDLINE=98409296; PubMed=9738916;  
 RA Maruyama N., Ichise K., Katsube T., Kishimoto T., Kawase S.,  
 RA Matsumura Y., Takeuchi Y., Sawada T., Utsuni S.;  
 RT Identification of major wheat allergens by means of the Escherichia  
 RT coli expression system.";  
 RL Eur. J. Biochem. 255:733-745(1998).  
 DR EMBL; D84341; BAI2318.1; - "reservoir activity; IEA."  
 DR GO; GO:0045735; F:nutrient reservoir activity;  
 DR InterPro; IPR003612; AAI.  
 DR InterPro; IPR001376; Gliadin.  
 DR InterPro; IPR001954; Gliadin.  
 DR Pfam; PF00234; Tryp\_alpha\_amy1; 1.  
 DR PRINTS; PR00208; GLTADSLTEN.  
 DR PRINTS; PR00209; GLIADIN.  
 DR SMART; SM00499; AAI; 1.  
 DR NON TER 1  
 FT CHAIN <1 259 alpha-gliadin mature peptide.  
 SO SEQUENCE 259 AA; 29996 MW; FE36CD48FD8F54C6 CRC64;  
 Query Match 91.9%; Score 1315.5; DB 2; Length 259;  
 Best Local Similarity 93.2%; Pred. No.1.8e-72;  
 Matches 248; Conservative 2; Mismatches 9; Indels 7; Gaps 1  
 QY 1 VRPVVQLPQNPSPQQQEQEQEYPLVQDQDQFPFGQDQFPFPQPYFPQPFPSGQPYQLQDP 60  
 DB 1 VRPVVQLQDQNPSPQQQPEQGVPLVQDQDQLGQDQFPFPQDQPYFPQPFPSGQPYQLQDP 60  
 QY 61 FPPQLPYPQPSFPFQPYFPFQPYQYQFPQDPTISQDQAQDQDQDQDQDQDQDQDQDQDQ 120  
 DB 61 FPPQLPYPQPSFPFQPYFPFQPYQYQFPQDPTISQDQAQDQDQDQDQDQDQDQDQDQDQ 113  
 QY 61 FPPQLPYPQPSFPFQPYFPFQPYQYQFPQDPTISQDQAQDQDQDQDQDQDQDQDQDQDQ 113  
 DB 121 QQLIPCDVVLQCHNIAHRSQVLDQSTYQLLELCCQHLWQIPDSQCAIHNVVAII 180  
 DB 114 QQLIPCDVVLQCHNIAHRSQVLDQSTYQLLELCCQHLWQIPDSQCAIHNVVAII 173  
 QY 181 LHQDQKQDQPSQSVFQQLQCYFLGGGSRFSSQDNPQDAGSYQFQQLQPFERITLAL 240  
 DB 174 LHQDQKQDQPSQSVFQQLQCYFLGGGSRFSSQDNPQDAGSYQFQQLQPFERITLAL 233  
 QY 241 QTLPAMCNVYIAPYCTIAPEGIFGNT 266  
 DB 234 QTLPAMCNVYIAPYCTIAPEGIFGNT 259  
 RESULT 14  
 Q41530 PRELIMINARY; PRT; 288 AA.  
 AC Q41530;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DE 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 OS Alpha-gliadin storage protein.  
 OC Triticum aestivum (Wheat).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
 OC Triticaceae; Triticum.  
 OC NCBI\_TaxID=4565;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RP STRAIN=Cheyenne;  
 RA Anderson O.D.;  
 RA Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.



DR EMEL, U51304, AAA96523.1, -.  
DR PIR, S13333, S13333.  
DR PIR, T06282, T06282.  
DR GO, GO:0045735, F:nutrient reservoir activity; IEA.  
DR InterPro, IPR003612, AAI.  
DR InterPro, IPR001376, Glia\_din.  
DR InterPro, IPR001954, Glia\_gluenin.  
DR Pfam, PF00234, Tryp\_alpha\_amy1, 1.  
DR PRINTS, PR00208, GLIADSLUTEN.  
DR PRINTS, PR00209, GLIADIN.  
DR SMART, SM00439, AAI, 1.  
DR SEQUENCE, 288 AA; 3385 MW; D7F6B99131283CA2 CRC64;

Query Match	91.7%	Score 1312;	DB 2;	Length 288;
Best Local Similarity	92.5%;	Pred. No. 3.2e-72;		
Matches 248;	Conservative 4;	Mismatches 14;	Indels 2;	Gaps 1;

QY	1	VRVVVPQLOLPONBPSQQQOPQEOVLVLVQQQQPFGQQQQFFPQQPYHQPPFFSQQPYILOLP	60
Db	21	VKRVFVQLOLPONBPSQQQOPQEOVLVLVQQQQPFGQQQQFFPQQPYHQPPFFSQQPYILOLP	80
QY	61	FPQRLPYQPQGSFPQPCPYPOQPCYGSQPCQPIPSQQQAQQQQQQQQQQQQ--QQILQQI	118
Db	81	FSQPLPYQPQGFRRQCPYPPQPCYGSQPCQPIPSQQQQQQQQQQQQQQQQQQILQQI	140
QY	119	LQQQLPCMDVVLQCHNIHANSQVLYQGSTYQLQELCCQHLMOIPBQSQCOALHNVTM	178
Db	141	LQQQLPCMDVVLQCHNIRGRSQVLYQGSTYQLQELCCQHLMOIPKLCQCAIHNVTM	200
QY	179	IIHHQQQXQQQPSQSSQSFQCPLOQYPLQGGSFRPSQNPQAGQSVQPOOLPQFEELRNL	238
Db	201	IIHHQQQXQQQPSQSSQSFQCPLOQYPLQGGSFRPSQNPQAGQSVQPOOLPQFEELRNL	260
QY	239	ALQTLPMQCNVITAPYCTIAPFGIRGTN	266
Db	261	ALQTLPMQCNVITAPYCTIAPFGIRGTN	288

```

RESULT: 15
Q01528.1 PRELIMINARY; PRT; 287 AA.
ID Q01528
AC Q01528;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-MAR-2004 (TRENBLREL. 26, Last annotation update)
DE Alpha-glutadin.
OS
   Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Magnoliopsida; Liliopsida; Poales; Poaceae
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae
OC Triticeae; Triticum.
OX NCBI_Taxid=4565;
   [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Cheyenne;
RA Anderson O.D.
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RL EMBL; U50984; AA96276.1; -.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR003612; AAI.
DR InterPro; IPR001376; Glutadin.
DR InterPro; IPR001954; G1ta_glutenin.
DR Pfam; PF00234; TYP_alpha_aryl; 1.
DR PRINTS; PR00208; GLINDGUTEN.
DR PRINTS; PR00209; GLIDIN.
DR SMART; SMO0297; AAI; 1.
SQ SEQUENCE 287 AA; 3319 MW; 05F82256749C9E97 CRC64;

```

```

Db      21 VREPVQLOPNPSQQLPEOEYVVLVQQQQQFGQQQPPRPQOPRYQPQ-PFSQLPYVQLQF 79
QY      61 FPGPRLPYPOPSFPBPQPYPPQPODQYSQPQOPIS-QQQAQQQQQQQQQQQQQQLLQQL 119
      120 QQQQLPCMVVLCQNNHIAHARSQVQSTYQLLDELCCQHLMOJPEOSQCAHNVVHAI 179
Db      80 FPGPQLPYQGPQFPBPQOPYPQGPQVQYQBPQOPISQCCQQQQQQQQQQQQQQQLLQQL 139
QY      120 QQQQLPCMVVLCQNNHIAHARSQVQSTYQLLDELCCQHLMOJPEOSQCAHNVVHAI 179
      140 QQQQLPCMVVLCQNNKAGRSQVQSTYQLLDELCCQHLMOJPEOSQCAHNVVHAI 139
Db      180 ILHQQX-QQQQPSQVSFQCPLOQYPLGGQSFRRSQQNPQAQSVQPPQLPQFEETRNL 238
QY      200 ILHQQQKQQQQQPSQSVSQCPLOQYPLGGQSFRRSQQNPQTSQSVQPPQLPQFEETRNL 259
Db      239 ALQTLPMQNVYIAPYCTIAPFGIRGTN 266
QY      260 ALQTLPSMCNVYIIPYCTIAPFGIRGTN 287

```

Query Match	90.8%;	Score 1299.5;	DB 2;	Length 287;
Best Local Similarity	92.9%;	Pred. No. 1.8e-71;		
Matches 249; Conservative	3;	Mismatches 13;	Indels 3;	Gaps 3

QY 1 VRVPQLQPCNPSQQQPEQVVLVQQQCFGQQQCFPPQPPQPPSPQQPYLQLP 60



This Page Blank (uspto)







The present invention describes a method (M1) for preventing or treating

[illegible]

DR WPI; 2003-697466/66







XX The present sequence represents the mature glutenin protein. The DNA  
 CC sequence encoding this protein is isolated from the genomic DNA of  
 CC Triticum durum 11. The gene codes for a low-molecular-weight glutenin  
 CC protein and can be used to produce transgenic durum wheat plants with  
 CC "better quality characteristics" (no details given). (Updated on 25-MAR-  
 CC 2003 to correct PI field.) (Updated on 17-OCT-2003 to standardise OS  
 CC field)

XX Sequence 369 AA;

Query Match 32.3%; Score 464; DB 2; Length 369;  
 Best Local Similarity 40.2%; Pred. No. 9.8e-32;  
 Matches 134; Conservative 32; Mismatches 87; Indels 80; Gaps 14;

6 FOLQFQNFSSQDP--OEQVPLVQDQPFSSQDQPF--PQDPYDPQPP----- 50  
 38 FQDQFCSQDQDQPPFSSQDQDQPPFSSQDQDQPPFSSQDQDQPPFSSQDQDQPP 97  
 51 ---SQDPLV-----QLQFQFQFQFQFQFQFQFQFQFQFQFQFQFQFQFQFQF 89  
 98 FSSQDQ 157  
 90 PQDQFSSQDQ 132  
 158 QDQFQ 216  
 133 CH-----NIAHARSQVLQSTYQLLQELCCQHTMQIPESQCCALHNVAHAIILHQDQK 186  
 217 CQCSFMAMPQSLARSQMTQSSSCHVMQDCCQDLPQIPQDSRYEALIVYSIIL--OEQ 274  
 187 CQDQ 234  
 275 QDQ 331  
 235 IENLALQTLPAKCNVYIAPV--CTTAPFGIFGT 265  
 332 MTSIALRTLPKCNWNPVLYRTTETVPPGV--GT 363

RESULT 6  
 ADO71669

ID ADO71669 standard; protein; 297 AA.

AC ADO71669;

DT 12-AUG-2004 (first entry)

XX Amino acid sequence of a modified glutenin LMW subunit.

KM low molecular weight subunit; LMW subunit; glutenin;  
 KW wheat cultivar Cheyenne; gliadin; flour; tablet; coeliac disease;  
 XX gluten intolerance.

XX Triticum sp.

OS Synthetic.

XX EPI424342-A1.

XX 02-JUN-2004.

XX 27-NOV-2002; 2002EP-00026461.

XX 27-NOV-2002; 2002EP-00026461.

XX (BAKE-) BAKEMARK DEUT GMBH.

XX (MONS) MONSANTO AGRAR DEUT GMBH.

XX (TUNF-) TUNFERN GMBH & CO KG.

XX (PURA-) PURATOS NV.

XX Hinzmann E, Wieser H, Stahl U;  
 XX WPI; 2004-402870/38.

DR N-PSDB; ADO71668.  
 XX Novel nucleic acid comprising sequence encoding modified glutenin  
 PT polypeptide, useful for preparing modified glutenin polypeptide as  
 PT gliadin substitute in foodstuffs such as dough, pastries and wafers.  
 XX Claim 16; Fig 11; 43pp; English.

XX The present sequence represents a modified low molecular weight (LMW)  
 CC subunit of glutenin. The wild type subunit is designated clone LMW6, and  
 CC is isolated from wheat cultivar Cheyenne. The LMW6 polypeptide does not  
 CC contain the allergenic epitope QQQPP, and shows some minor differences to  
 CC published sequences. It therefore represents a new allele for LMW subunit  
 CC genes. The LMW6 polypeptide was modified to produce modified glutenin  
 CC polypeptides of the invention. In these modified polypeptides one or more  
 CC cysteine residues responsible for intermolecular cross linking through  
 CC disulfide bridges are deleted or substituted. It is also useful in the  
 CC polypeptide is useful as a gliadin substitute. It is also useful in the  
 CC preparation of foodstuffs, such as flour or for the preparation of  
 CC pharmaceutical products, such as tablets, where the foodstuffs contain a  
 CC considerably reduced amount of gliadin proteins or no gliadin proteins.  
 CC pharmaceutical compositions comprising the modified polypeptide of the  
 CC invention are useful for treating patients suffering from coeliac disease  
 CC or persons who are intolerant to gluten.

XX Sequence 297 AA;

Query Match 31.4%; Score 450.5; DB 8; Length 297;  
 Best Local Similarity 41.2%; Pred. No. 1.1e-30;  
 Matches 120; Conservative 38; Mismatches 82; Indels 51; Gaps 13;

5 VPQLQFQNFSSQDP--OEQVPLVQDQPFSSQDQDQPFSSQDQDQPFSSQDQDQPFSSQ 63  
 18 IQWERTSIFGLRPMQDQ 71  
 64 PWLPRYDQ 116  
 72 -----QPLFSQKQDQ 125  
 117 QI-----LQDQLIPCMQDQ 170  
 126 QLNPKCVPLQDQ---CSPVAMPQH---LARSQMWQSSCNWQDCCQQLPRIPQSRYE 179  
 171 AIHNVAHAIILHQDQK-----QQQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQ 213  
 180 AIRAIRIFSIILQEQDQ 239  
 214 SQDQ 262  
 240 QDQ 290

RESULT 7  
 ADO71661

ID ADO71661 standard; protein; 298 AA.

AC ADO71661;

DT 12-AUG-2004 (first entry)

XX Amino acid sequence of glutenin clone LMW6.

KM low molecular weight subunit; LMW subunit; glutenin;  
 KW wheat cultivar Cheyenne; gliadin; flour; tablet; coeliac disease;  
 XX gluten intolerance.

XX Triticum sp.

OS Location/Qualifiers

FT Key Disulfide-bond 25 /note = this residue forms an intermolecular disulfide  
 FT bond  
 FT Disulfide-bond 227







CC No interference between the individual RNA sequences occur. This sequence  
CC represents a protein encoded by a target gene used in the method of the  
CC invention.

CC Sequence 307 AA;

Query Match 30.4%; Score 437.5; DB 7; Length 307;  
Best Local Similarity 42.8%; Pred. No. 1.5e-29;  
Matches 125; Conservative 36; Mismatches 76; Indels 55; Gaps 16;

CC 13 PSQQQPEQGVPLVQQQCFP-----GQQQQPFPQGPYPQGPFPFQQQPYLQQLPFPQWLP 67  
CC 27 PGLERPMWQQQLPFPQGTFPQPLFSQQQ-----QLFPQGPFSQQQP-----PFWQQQPP 78  
CC 68 YPQPSFPQGP-----YPQPSQPSQGPPI-----SQQQAQQQQQQQQQQQQQQQ1 114  
CC 79 FSQQQPLPFPQGPFSQQQQLVLPQ-QPFPFSQQQQLVLPQGPFSQQQQLVQQQIP 137  
CC 115 -LQQLIQQLIPQMDVVLQCH--NIH---ARQVLTQSTYQLQLQELCCQHLMQIPRQS 167  
CC 138 VVQPSILQQLNQC-KVFLQCCSPVAMPQRLARSQMLQSSCHWQQCCQQLPQIPQS 196  
CC 168 QCAIHNVVHAILLHQQKQKQSSQVSPQPLQGYPLGGSFRPSQWNPQ-----219  
CC 197 RYEAIRAITIYSTILL-QEQQVQSGISQSQQQPQQ---LGCVSGPQQSQSQQLGQQPQQ 251  
CC 220 ---AGGS-VQPSQLPQFEIRNLALQTLPAWCVYIAPY-CTIAPGIFGT 265  
CC 252 QQLAQGTFLQPHQIAQLEWMTSIALRLIPTWCNVVPLXRTTTSVPRGV-GT 302

RESULT 9

ADG44134  
ADG44134 standard; protein; 307 AA.

ADG44134;

26-FEB-2004 (first entry)

T. aestivum glutenin-1D1 protein.

oil content; plant; storage protein; seed-specific promoter; 2S-albumin;  
7S-globulin; 11S-globulin; 12S-globulin; zein-prolamine; transgenic;  
oil production; fat production; free fatty acid production; food;  
animal feed; pharmaceutical; fine chemical production; glutenin.

Triticum aestivum.

WO2003077643-A2.

25-SEP-2003.

17-MAR-2003; 2003WO-EP002733.

20-MAR-2002; 2002DE-01012893.

(BADI ) BASF PLANT SCI GMBH.

Bauer J;

WPI; 2004-011485/01.

N-PSDB; ADG44133.

Increasing total oil content of plants, useful e.g. as foods or animal  
feeds, by reducing amount of storage proteins, particularly with double-  
stranded interfering RNA.

Claim 4; SEQ ID NO 174; 253bp; German.

This invention describes a novel method for increasing the total oil  
content of a plant by reducing the amount of at least one storage protein  
in the plant (or its tissue, organs, parts or cells) and selecting plants  
that have higher total oil content than starting plants. The storage

CC protein is suppressed by introducing antisense RNA, optionally combined  
CC with a ribozyme, sense RNA that induces co-suppression, DNA-binding  
CC factors directed against storage protein genes, viral sequences that  
CC degrade storage protein RNA, constructs that induce homologous  
CC recombination of endogenous storage protein genes or mutations into  
CC storage protein genes. Most preferably a plant cell is stably transfected  
CC with a recombinant expression construct, then regenerated to plants that  
CC express the incorporated sequence. The expression constructs particularly  
CC contain a seed-specific promoter and they are introduced into plants by  
CC standard methods, e.g. via Agrobacterium. The preferred storage proteins  
CC of the invention are 2S-albumins, 7S or 11S/12S-globulins or zein-  
CC prolamines. Transgenic organisms produced by the new method are used for  
CC production of oils, fats, free fatty acids or their derivatives, useful  
CC as foods, animal feeds, pharmaceuticals and fine chemicals. This sequence  
CC represents a storage protein used to illustrate the method of the  
CC invention.

CC Sequence 307 AA;

Query Match 30.4%; Score 437.5; DB 8; Length 307;  
Best Local Similarity 42.8%; Pred. No. 1.5e-29;  
Matches 125; Conservative 36; Mismatches 76; Indels 55; Gaps 16;

CC 13 PSQQQPEQGVPLVQQQCFP-----GQQQQPFPQGPYPQGPFPFQQQPYLQQLPFPQWLP 67  
CC 27 PGLERPMWQQQLPFPQGTFPQPLFSQQQ-----QLFPQGPFSQQQP-----PFWQQQPP 78  
CC 68 YPQPSFPQGP-----YPQPSQPSQGPPI-----SQQQAQQQQQQQQQQQQQQ1 114  
CC 79 FSQQQPLPFPQGPFSQQQQLVLPQ-QPFPFSQQQQLVLPQGPFSQQQQLVQQQIP 137  
CC 115 -LQQLIQQLIPQMDVVLQCH--NIH---ARQVLTQSTYQLQLQELCCQHLMQIPRQS 167  
CC 138 VVQPSILQQLNQC-KVFLQCCSPVAMPQRLARSQMLQSSCHWQQCCQQLPQIPQS 196  
CC 168 QCAIHNVVHAILLHQQKQKQSSQVSPQPLQGYPLGGSFRPSQWNPQ-----219  
CC 197 RYEAIRAITIYSTILL-QEQQVQSGISQSQQQPQQ---LGCVSGPQQSQSQQLGQQPQQ 251  
CC 220 ---AGGS-VQPSQLPQFEIRNLALQTLPAWCVYIAPY-CTIAPGIFGT 265  
CC 252 QQLAQGTFLQPHQIAQLEWMTSIALRLIPTWCNVVPLXRTTTSVPRGV-GT 302

RESULT 10

ABB71695  
ABB71695 standard; protein; 1798 AA.

ABB71695;

26-MAR-2002 (first entry)

Drosophila melanogaster polypeptide SEQ ID NO 41877.

Drosophila; developmental biology; cell signalling; insecticide;

pharmaceutical.

Drosophila melanogaster.

WO200171042-A2.

27-SEP-2001.

23-MAR-2001; 2001WO-US009231.  
23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.

(PEKE ) PE CORP NY.

Venter JC, Adams M, Li PWD, Myers EW;

WPI; 2001-656860/75.



[illegible]

PS	Disclosure; SEQ ID NO 15963; 21pp + Sequence Listing; English.
XX	
CC	The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB116176-AB10511), expressed DNA sequences (AB101840-AB116175) and the encoded proteins (ABBS7737-ABBS7927). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX	
Sequence 2265 AA:	
Query Match	18.6%; Score 268; DB 4; Length 2285;
Best Local Similarity	35.7%; Pred. No. 6.6e-14;
Matches 101; Conservative 14; Mismatches 104; Indels 64; Gaps 13;	
Dy	7 QLOPFNP-----SQGQPOEQLVLYGGQQPFGQQQQPPQPQFPFPOPQ----PPF-SQQP 54
Dd	909 QMGQQQPLVFMAFTYNHGVMPQQVN--QQQQQPMQMGIPLQGVYQDQDPVTLPFRPHBQPF 966
Dy	55 YLQLQPFPO-----FWLPYPPOGSFPQOPYPPOBPQYSQPQPISSQQA--- 99
Dd	967 QQQQQQPLQQLQGLMTHTNVGAPEL--TQQQQMAQQQAQQVYFQQQQQQCPQQAIVNMQAAAYAM 1024
Dy	100 CQQQQQQQQQQQQCQIILQCIILQQCL-----PCMDVYLQCNINAHASQVLQGSTYQLQ 153
Dd	1025 QQAQQQQQLSPQLDIQQCIILQQQQVAVSHQQQIMQQQLAQHQQLQQLQQQLQQQQQLQQQQ 1084
Dy	154 ELCCQHLMQEPESQCAIHNVHAIIHLQQKQKQQQQPSSQV---SFQGEPLQ-----QY 204
Dd	1085 QIQQQQLLQQQQLQQQ-QVGCGYAGA--MPPQGHHQLWTGSQVMAFHNGQQGTQLIPVMQVQ 1144
Dy	205 PL-----GGGSFPSQQNPDAQSGSYVPQQLPQF 232
Dd	1142 PPTSVAPPIQHTYNQGGGYVTLSDAQQQMHPRGFSAVPOQAAPF 1184
RESULT 12	
ID	AA96255 standard; protein, 1162 AA.
AA96255	
AC	AA96255;
XX	
DT	12-SEP-2003 (revised)
DT	11-SEP-2000 (first entry)
XX	
DE	Kaposi's sarcoma-associated herpesvirus LANA.
KW	Kaposi's sarcoma-associated herpesvirus; KSHV; rhadino virus;
KW	latency-associated nuclear antigen; LANA; gamma-2 herpes virus;
KW	human herpes virus 8; HHV8; rhadino virus cis-acting element; RVCAE;
KW	Kaposi's sarcoma; primary effusion lymphoma; PEL;
KW	human immunodeficiency virus; HIV; multicentric Castleman's disease.
OS	Human herpesvirus 8.
XX	
Key	Location/Qualifiers
FH	14..17
FT	/note= "nuclear localisation signal, NLS"
FT	64..70
FT	/note= "nuclear localisation signal, NLS"
FT	320..429
FT	/note= "acidic repeat region"
FT	430..549
FT	/note= "Gln, Glu, Pro-rich region"
FT	550..589
FT	Region



D6 495 PLQEPQQQEPPQQQEPPQQQEPPQQQEPPQQQEPPQQQEPPQQQE-  
FQQQEFPNN











GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compen Ltd.

## OM protein - protein search, using sw model

Run on: December 14, 2004, 17:06:20 ; Search time 19 Seconds

(without alignments)  
928.452 Million cell updates/sec

Title: US-10-089-700-3-w65

Perfect score: 1437

Sequence: 1 VRRVPRQLQPNPSQQQPE.....CNVYIAFYCTIAFPFGFTGN 266

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 6631800 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/prodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/prodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/prodata/1/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/1/iaa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	464	32.3	369	2	US-08-991-300-2
2	265	18.4	1162	2	US-08-728-323A-2
3	265	18.4	1162	3	US-09-298-568-2
4	265	18.4	1162	4	US-09-410-399-2
5	265	18.4	1162	4	US-09-894-273-2
6	247	17.2	788	2	US-08-918-914-4
7	236	16.4	498	4	US-09-270-767-45042
8	234	16.3	256	4	US-09-248-796A-21251
9	215.5	15.0	579	4	US-09-668-119-3
10	215	15.0	2074	4	US-09-491-356C-9
11	201	14.0	2023	4	US-09-491-356C-8
12	201	14.0	2124	4	US-09-538-092-1377
13	198.5	13.8	505	4	US-09-248-796A-19253
14	197.5	13.7	663	4	US-09-270-767-61220
15	197.5	13.7	1591	4	US-09-270-767-45698
16	197.5	13.7	2441	1	US-08-194-468-2
17	197.5	13.7	2441	3	US-08-961-739-2
18	197.5	13.7	2441	3	US-09-514-247A-8
19	197.5	13.7	2441	3	US-09-686-316-2
20	196.5	13.7	379	4	US-09-248-796A-23759
21	196.5	13.7	2442	4	US-09-514-247A-10
22	196.5	13.7	2442	4	US-09-538-092-1370
23	196	13.6	216	4	US-09-248-796A-21017
24	194.5	13.5	729	3	US-09-625-188-20
25	185	12.9	295	4	US-09-248-796A-20004
26	184.5	12.8	316	4	US-09-270-767-42663
27	184.5	12.8	320	4	US-09-248-796A-24758

28	184.5	12.8	332	4	US-09-248-796A-21649	Sequence 21649, A
29	183	12.7	169	4	US-09-248-796A-28087	Sequence 28087, A
30	183	12.7	519	4	US-09-248-796A-19263	Sequence 19263, A
31	178	12.4	383	4	US-09-248-796A-23236	Sequence 23236, A
32	178	12.4	408	4	US-09-248-796A-14480	Sequence 14480, A
33	178	12.4	1319	4	US-09-538-092-1291	Sequence 1291, Ap
34	177	12.3	261	4	US-09-602-565-34	Sequence 34, Appl
35	176.5	12.3	675	4	US-09-248-796A-20659	Sequence 20659, A
36	176.5	12.3	684	4	US-09-823-240A-9	Sequence 9, Appl
37	176	12.2	618	4	US-09-248-796A-15319	Sequence 15319, A
38	176	12.2	848	4	US-09-538-092-33	Sequence 33, Appl
39	174.5	12.1	382	4	US-09-248-796A-18720	Sequence 18720, A
40	174	12.1	657	4	US-09-248-796A-19232	Sequence 19232, A
41	173	12.0	1507	4	US-09-914-359-37	Sequence 37, Appl
42	171.5	11.9	903	2	US-08-853-310-2	Sequence 2, Appl
43	168.5	11.7	311	4	US-09-248-796A-27827	Sequence 27827, A
44	167.5	11.7	667	2	US-08-718-661-2	Sequence 2, Appl
45	167	11.6	542	1	US-07-814-964-13	Sequence 13, Appl

## ALIGNMENTS

RESULT 1  
US-08-991-300-2  
Sequence 2, Application US/08991300  
Patent No. 5973225  
GENERAL INFORMATION:  
APPLICANT: D/OVIDIO, RENATO  
APPLICANT: PORCEDDU, ENRICO  
APPLICANT: MERCHITELLI, CINZIA  
APPLICANT: CARDELLI, LUISA ERCOLI  
TITLE OF INVENTION: ISOLATION AND CHARACTERIZATION OF A GENE  
TITLE OF INVENTION: ENCODING A LOW MOLECULAR WEIGHT GLUTENIN  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSER: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/991.300  
FILING DATE: 16-DEC-1997  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: IT MI 96/A 002663  
FILING DATE: 19-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 2264-0201-0X  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
FAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 369 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-991-300-2  
Query Match 32.3%; Score 464; DB 2; Length 369;  
Best Local Similarity 40.2%; Pred. No. 2.3e-35;



Matches 134; Conservative 32; Mismatches 87; Indels 80; Gaps 14;

QY 6 PQLQPNFSSQDP-----QEQVPLVQGGFPFGQGGQGF--PQQPFPQGPFP-----50  
 DB 38 PQQGPGCSQQQQQPLSSQQQPPFSQQQPPVLPQQPFSQQQLPPFSQQQCP 97  
 QY 51 ---SQQPYL-----QLQFPQPM-LPYQPQSPFPQ-----PYQPQPPQYSQ 89  
 DB 98 PQQQQQPVLPQQPFSQQQLPPFSQQQLPPFSQQQLPPFSQQQLPPFSQQ 157  
 QY 90 PQQPISQQQAQQQQQQQ-----QQQQQQLLQ-----LLQQQLPPQMDVYLQ 132  
 DB 158 QQQPVLPPQQPPFSQQQPPFPQPPFSQQQPPVLLQQQLPFPVPSILQLPNC-KVFLQ 216  
 QY 133 QH-----NIHARSQVLSQSTYQLLQELCCCHLMQIPESQCAIHNVVHAILHQQX 186  
 DB 217 QQQSWAMPQSLARQMLQSSCHVMQCCQCCQLPQIPQGRVRAIRAIYVSIIL--QEQ 274  
 QY 187 QQQQSSQSVSFQQLQGYPLQGGSFPSQGNPQAQGS-----VQPPQLPQFEE 234  
 DB 275 QGVQGSIGVQQQQPQ--LQCVSQPQCGSQQLGQQPQQQLAHGTFLOPHQIAQLLEV 331  
 QY 235 IENALQTLPMQCNVYIAPY--CTIAPGIGT 265  
 DB 332 MTSIALRTLPQCNMVPVLTTRTVPPGV-GT 363

## RESULT 2

US-08-728-323A-2  
Sequence 2, Application US/08728323A

Patent No. 5948676

GENERAL INFORMATION:

APPLICANT: Chang, Yuan

APPLICANT: Bohenzky, Roy A.

APPLICANT: Russo, James V.

APPLICANT: Edelman, Isidore S.

APPLICANT: Moore, Patrick S.

TITLE OF INVENTION: Immediate Early Protein From Kaposi's

TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA

TITLE OF INVENTION: Encoding Same And Uses Thereof

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESS: Cooper &amp; Dunham LLP

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REFERENCE NUMBER: 28,678

TELEPHONE: 212-278-0400

TELEFAX: 212-391-0525

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1162 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

Query Match

18.4%; Score 265; DB 2; Length 1162;

Best Local Similarity 38.0%; Pred. No. 3,1e-16;  
Matches 93; Conservative 19; Mismatches 101; Indels 32; Gaps 10;

QY 4 PVPQLQPNFSSQDPQEQVPLVQ--QQQPPGQ--QFPFPQPPFPQPPFSQPPYLQ 59  
 DB 495 PQQPQDPQDPQDPQDPQDPQDPQDPQDPQDPQDPQDPQDPQDPQDPQDP 553  
 QY 60 P--FPQ--PWLFPYQ--PQSPFPQPPYPP--QPVYSPQPPISQQAQQQQQQQQ 109  
 DB 554 PQQREPQQRFPQQRFPQQRFPQQRFPQQRFPQQRFPQQRFPQQRFPQQR 613  
 QY 110 QQQQLLQQLLQQLLPCMDVVLQGNHARSQVLSQSTYQLLQELCCCHLMQIPESQ 169  
 DB 614 EQQDEQQDEQQ--QEQQDEQQDEQQDEQQDEQQDEQQDE--QQQDEQQQD 663  
 QY 170 QAIHNVHAILHQQQKQQQQSSQSVSFQQLQGYPLQGGSFPSQGNPQAQSVQPOL 229  
 DB 664 EQQD-----EQQDDEQQDEQQDEQQDEQQDEQQDEQQDEQQDEQQDE 716  
 QY 230 PQFEE 234  
 DB 717 EQQD 721

## RESULT 3

US-09-298-568-2  
Sequence 2, Application US/09298568

Patent No. 632792

GENERAL INFORMATION:

APPLICANT: Kleff, Elliott D.

APPLICANT: Balleslas, Mary E.

APPLICANT: Kaye, Kenneth M.

TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO

TITLE OF INVENTION: VIRUS DNA TO MEDIATE EFFICIENT EPISOME PERSISTENCE

FILE REFERENCE: 16412-10001R

CURRENT APPLICATION NUMBER: US/09/298,568

EARLIER FILING DATE: 1999-04-21

EARLIER APPLICATION NUMBER: US 60/109,422

EARLIER FILING DATE: 1998-11-19

NUMBER OF SEQ ID NOS: 3

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 2

LENGTH: 1162

TYPE: PRT

ORGANISM: Kaposi's sarcoma-associated herpesvirus

US-09-298-568-2

Query Match

18.4%; Score 265; DB 3; Length 1162;  
Best Local Similarity 38.0%; Pred. No. 3,1e-16;  
Matches 93; Conservative 19; Mismatches 101; Indels 32; Gaps 10;

QY 4 PVPQLQPNFSSQDPQEQVPLVQ--QQQPPGQ--QFPFPQPPFPQPPFSQPPYLQ 59  
 DB 495 PQQPQDPQDPQDPQDPQDPQDPQDPQDPQDPQDPQDPQDPQDPQDPQDP 553  
 QY 60 P--FPQ--PWLFPYQ--PQSPFPQPPYPP--QPVYSPQPPISQQAQQQQQQQQ 109  
 DB 554 PQQREPQQRFPQQRFPQQRFPQQRFPQQRFPQQRFPQQRFPQQRFPQQR 613  
 QY 110 QQQQLLQQLLQQLLPCMDVVLQGNHARSQVLSQSTYQLLQELCCCHLMQIPESQ 169  
 DB 614 EQQDEQQDEQQ--QEQQDEQQDEQQDEQQDEQQDEQQDE--QQQDEQQQD 663  
 QY 170 QAIHNVHAILHQQQKQQQQSSQSVSFQQLQGYPLQGGSFPSQGNPQAQSVQPOL 229  
 DB 664 EQQD-----EQQDDEQQDEQQDEQQDEQQDEQQDEQQDEQQDEQQDE 716  
 QY 230 PQFEE 234  
 DB 717 EQQD 721

## RESULT 4















APPLICANT: Keith Weinstein et al  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
 TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
 FILE REFERENCE: 1571196.132  
 CURRENT APPLICATION NUMBER: US/09/248,796A  
 CURRENT FILING DATE: 1999-02-12  
 PRIOR APPLICATION NUMBER: US 60/074,725  
 PRIOR FILING DATE: 1998-02-13  
 PRIOR APPLICATION NUMBER: US 60/096,409  
 PRIOR FILING DATE: 1998-08-13  
 NUMBER OF SEQ ID NOS: 28208  
 SEQ ID NO 19253

Query Match	13.8%;	Score 138.5;	DB 4;	Length 505;
Best Local Similarity	27.1%;	Pred. No. 1.7e-10;		
Matches 64;	Conservative 24;	Mismatches 47;	Indels 10;	Gaps 7.

[illegible]

RESULT 14  
 US-09-270-767-61220  
 Sequence 61220, Application US/09270767  
 Patent No. 6703491  
 GENERAL INFORMATION:  
 APPLICANT: Homdurger et al.  
 TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*  
 PUBL REFERENCE: File Reference: 7326-054  
 CURRENT APPLICATION NUMBER: US/09/270,767  
 CURRENT FILING DATE: 1999-03-17  
 NUMBER OF SEQ ID NOS: 62517

Query Match	13.7%	Score 197.5;	DB 4;	Length 663;
Best Local Similarity	29.8%	Pred. No. 3e-10;		
Matches	88;	Conservative	25;	Mismatches 97;
				Indels 85;
				Gaps 13

QY	4	PVTOI.QONPSSQOQOEQOEVPLVVOOQOQFPGQOQOQSPFGQOQPYFQOQOQPPSSQOQPYLOLOPFQO	63
Db	195	PVAE---BEOBEOQAEHBOYAVS:QOPQOQLO:-----QGRPQYGGPY-QL-PLP-	242
QY	64	PMLPYPOPSPPQOPPYPOPOQYSSQOPQISQOQ-----AQOQOQO	105
Db	243	--LPAGHRSVNPQ-----QOQOQOHQPOVPIPEBQFLKLEBLOARAYHEQLRQOQOQ	296
QY	106	QOQOQOQOQLOQLOILQOQOILPCMDVYLQOHNINAHRSQVLOQSTYOL-----LOE	154
Db	297	QOQOQOQOQOQOQOQH-----HKOLPHSHR-ATHKVLQOQADPISGLGGYBRPVAEBE	350
QY	155	LC-----COHLMQIPESQCCALHNHVHAILHQOQKQOQOQSSQYVS	196
Db	351	LWPTYSHRGGRKXLP.LPQAQOIQOEDEQOQOQOPQOVLHKPIPHPQOQLHGS.LPQITA	410
QY	197	FOQPLQOQYLLGQSGFRPQOQNPQAGSVOPOQLPQFEIENLALQTLPMACNYI	251
Db	411	YQP-----QISTYTLPHNP.LAKSSIE-----SEIKLILANKPQOSLAV	451

RESULT 15  
US-09-270-767-45698

```

; Sequence 43836; Application 09/270.767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270.767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ. ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45698
; LENGTH: 1591
; TYPE: PR1
; ORGANISM: Drosophila melanogaster
; US-09-270-767-45698

```

Query Match	13.7%;	Score 197.5;	DB 4;	Length 1591;
Best Local Similarity	29.8%;	Pred. No. 8.6e-10;		
Matches	88;	Conservative	25;	Mismatches 97;
				Indels 85;
				Gaps 13

```

QY 4 LVPLQAPBNBQQQQQQEQVPLVQQQQQFQQQQQQQQPPQQPVPBPQBPFFQQGVLQQLPPQQ 63
Db 1123 FVAE---BEQFQEDQAEQVAVSTQBPQQQLQ-----QQPAPROYQGFY-QL-PLP- 11770
QY 64 PWLPYQBPQSPFPQQPVPBPQBPQVXSQBPQSPISQQQ-----AQQQQQQQ 105
Db 1171 --LPBPQRSVNPQ---QQQQQHQQPQVPIPEEQFLKILEEBYQARAYHQLQQQQQHQ 1224
QY 106 QQQQQQQQLLQQLLQQQLIPGMDVLLQOHNLNHRSGVQLQSTQQL-----LQE 154
Db 1225 QQQQQQQQHQQQQQQQH---HKQLPQHNST-NTKKVLQQAQDSLSLGGTRERFVAQE 1278
QY 155 LC-----CQHLWQIPBQSQAQHNHVAIILHQQKQKQQQSSQVS 196
Db 1279 LVTPVYSHRGGPKYLPFAQQLQIQEDDEQKQQQQQBPQVQLHKPQPHQSQLIHGLPQIA 1338
QY 197 FQCPQLQYPLDQGSFRPQQBPQAGSQVQRPQLLQFEEIRMLADQLPAMCNVYI 251
Db 1339 YQCP-----QISTKTLPNHPLKSSLE-----SIEETKLANKQGSLAYV 1379

```



Wed Dec 15 10:02:02 2004

us-10-089-700-3-w65.ra1

Page 7

Search completed: December(14, 2004, 17:29:02  
Job time : 20 secs

---



*This Page Blank (uspto)*



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 17:09:46 ; Search time 64.8333 Seconds  
(without alignments)  
1465.441 Million cell updates/sec

Title: US-10-089-700-3-w65

Perfect score: 1437

Sequence: 1 VEVFPVQLOPQNPSSQQQPOE.....CNVYIAPYCTIAPFIFGTN 266

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1585576 seqs, 357178320 residues

Total number of hits satisfying chosen parameters: 1585576

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB.pep:\*  
4: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep:\*  
5: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*  
6: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep:\*  
7: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubppaa/US11\_NEW\_PUB.pep:\*  
19: /cgn2\_6/ptodata/1/pubppaa/US11\_NEW\_PUB.pep:\*  
20: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1342.5	93.4	287	US-10-739-930-9777	Sequence 9777, Ap
2	1241.5	86.4	319	US-10-739-930-9619	Sequence 9619, Ap
3	1238	86.2	298	US-10-739-930-9770	Sequence 9770, Ap
4	587	40.8	327	US-10-739-930-9623	Sequence 9623, Ap
5	555.5	38.7	282	US-10-474-955-101	Sequence 101, App
6	544.5	37.9	298	US-10-739-930-9621	Sequence 9621, Ap
7	522	36.3	279	US-10-474-955-97	Sequence 97, Appl
8	521	36.3	279	US-10-474-955-99	Sequence 99, Appl
9	518	36.0	279	US-10-474-955-100	Sequence 98, Appl
10	518	36.0	279	US-10-474-955-100	Sequence 100, Appl
11	445.5	31.0	307	US-10-739-930-9778	Sequence 9778, Ap
12	437.5	30.4	307	US-10-739-930-9782	Sequence 9782, Ap
13	381	26.5	244	US-10-739-930-9769	Sequence 9769, Ap

14	277	19.3	541	US-10-425-115-200100	Sequence 200100,
15	274	19.1	283	US-10-425-115-200097	Sequence 200097,
16	265	18.4	1162	US-09-894-273-2	Sequence 2, Appl1
17	265	18.4	1162	US-10-294-804-2	Sequence 2, Appl1
18	263.5	18.3	1069	US-10-161-927-54	Sequence 54, Appl
19	254	17.7	323	US-10-425-114-44003	Sequence 44003, A
20	249	17.3	390	US-10-451-4678-64	Sequence 64, Appl
21	245.5	17.1	229	US-10-424-599-270450	Sequence 270450,
22	239.5	16.7	229	US-10-425-114-41056	Sequence 41056, A
23	238	16.6	358	US-10-104-047-3710	Sequence 3710, Ap
24	231.5	16.1	148	US-10-465-217-15	Sequence 15, Appl
25	228	15.9	1044	US-10-425-114-72709	Sequence 72709, A
26	226.5	15.8	362	US-10-425-115-261231	Sequence 261231,
27	226	15.7	738	US-10-451-4678-238	Sequence 238, App
28	225.5	15.7	192	US-10-425-114-52411	Sequence 52411, App
29	222	15.4	1173	US-10-437-963-144743	Sequence 144743,
30	221	15.4	1236	US-09-769-787-109	Sequence 109, App
31	219.5	15.3	1645	US-10-263-929-176	Sequence 176, App
32	219	15.2	4952	US-10-051-874-56	Sequence 56, Appl
33	219	15.2	5008	US-10-051-874-166	Sequence 166, App
34	219	15.2	5159	US-10-085-198-112	Sequence 112, App
35	219	15.2	5262	US-10-051-874-165	Sequence 165, App
36	219	15.2	5262	US-10-051-874-167	Sequence 167, App
37	216	15.0	351	US-10-264-049-2693	Sequence 129, App
38	215.5	15.0	1024	US-10-479-546-12	Sequence 12, Appl
39	215.5	15.0	1153	US-10-479-546-1	Sequence 1, Appl1
40	213.5	14.9	1351	US-10-282-1228-75147	Sequence 75147, A
41	213.5	14.9	1359	US-10-437-963-190032	Sequence 190032,
42	212.5	14.8	944	US-10-425-114-55600	Sequence 55600, A
43	211	14.7	1343	US-10-282-1228-75965	Sequence 75965, A
44	211	14.7	1362	US-09-815-242-14009	Sequence 14009, A
45	206.5	14.4	1025	US-10-437-963-113116	Sequence 113116,

## ALIGNMENTS

RESULT 1  
US-10-739-930-9777 Application US/10739930  
Sequence 9777, Publication No. US20040216390A1  
GENERAL INFORMATION: David K.  
APPLICANT: Kovalic, David K.  
TITLE OF INVENTION: NOCLETIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH  
FILE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT  
TITLE REFERENCE: 38-21(53377)B  
CURRENT APPLICATION NUMBER: US/10739, 930  
CURRENT FILING DATE: 2003-12-18  
NUMBER OF SEQ ID NOS: 11088  
SEQ ID NO 9777  
LENGTH: 287  
TYPE: PRT  
ORGANISM: Trifolium aestivum  
FEATURE:  
OTHER INFORMATION: Clone ID: TRIAE-23APR03-C176\_238.p  
US-10-739-930-9777

Query Match 93.4%; Score 1342.5; DB 17; Length 287;  
Best Local Similarity 95.1%; Pred. No. 5.2e-98;  
Matches 254; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

QY	1	VEVFPVQLOPQNPSSQQQPOE	US-10-739-930-9777	Sequence 9777, Ap
DB	21	VEVFPVQLOPQNPSSQQQPOE	US-10-739-930-9770	Sequence 9770, Ap
QY	61	FPQPMVFPQNPSSQQQPOE	US-10-474-955-101	Sequence 101, App
DB	81	FPQPMVFPQNPSSQQQPOE	US-10-474-955-97	Sequence 97, Appl
QY	121	OOIIPMDVVOOHNTIARSOVLOO	US-10-474-955-100	Sequence 98, Appl
DB	141	OOIIPMDVVOOHNTIARSOVLOO	US-10-474-955-100	Sequence 100, Appl







	Query Match	38.7%;	Score 555.5;	DB 17;	Length 282;	
	Best Local Similarity	47.6%;	Pred. No. 6.4e-36;			
	Matches 138;	Conservative 32;	Mismatches 73;	Indels 47;	Gaps 15;	
QY	2 RVPVPL-----QPQNPSQQQPFQEVLVQQQQPFGQ--QQGF-----PPQQYPFPD-QPRPS 51					
DG	10 QVMPFQQQPFDPQPFQSFQSQPQGTTP-QPGQTFFMQPQQQSOPQQPQQCPRIQDFQPPFP 68					
QY	52 --QQPYLQ-LQPFEPQWLP-YPOQSFPPQPFYQPOPFQYSOQFPQISQQAQQQQQQQ 106					
DG	69 PQGTTPGRQPPFFQQTQQPQGFQSQPQGFQPFQPPQGFQPFQGP-----QQS 118					
QY	107 QQQQQQQLQQQLIPCKMDVLTQHN----IHAASQYLQSTYLLAQELCCOHMQ 162					
DG	119 FPQQQPSLTIGSLQQQLNPKCNKFLTQQCKPVLSVSSMTLPSRDCGVMMRQCCQQLAQ 178					
QY	163 IPESQCCCAIHNVVAHIIHQQKQKQQQSSQVSFQCLP-QQYPLGGGSFRBSPQNPQAQ 221					
DG	179 IPQQIQCAAHISIVHSITIMQEQEQQRQ---GVQLTVLPSQQQVGQGTL-----VQGG 229					
QY	222 GSVQPDQLPQFEETRNLAQLTLPAMCNYIAFYCTI--APF-----GIKG 264					
DG	230 GIIDPQDAOLEVIRSLVLCITATACNCNVVPEPCSTIRAFPSIYAGIGG 279					

Query Match	37.9%	Score 544.5	DB 17	Length 298
Best Local Similarity	46.9%	Pred. No. 5e-35		
Matches 134	Conservative 28	Mismatches 75	Indels 49	Gaps 14
QY	4	PVPG:QPPNPQQQFEQVPLVQKQQFPQKQQQFP--PQPPYPPQPPFPSSQCPYVQLQDPF	61	

```

RESULT 7
US-10-474-955-97
; Sequence 97, Application US/10474955
; Publication No. US20040241161A1
; GENERAL INFORMATION:
; APPLICANT: Dr.Jihour, Jen W.
; APPLICANT: Koning, Frits
; APPLICANT: McAdam, Stephan N.
; APPLICANT: Ludvig, Sollid Magne
; TITLE OF INVENTION: METHODS AND MEANS FOR USE OF HLA-DQ RESTRICTED T-CELL RECEPTORS
; TITLE OF INVENTION: DO BINDING PROLAMINE-DERIVED PEPTIDES
; FILE REFERENCE: 2799/71244-PCT-US
; CURRENT APPLICATION NUMBER: US/10/474,955
; CURRENT FILING DATE: 2003-10-13
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 97
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
US-10-474-955-97

```

	Query Match	36.3%	Score 522	Length 279
	Best Local Similarity	46.2%	Pred. No. 2, 8e-33	
	Matches 132	Conservative 26	Mismatches 76	Incls 52; Gaps 14
QY	4 PVPQIQPNBSQQQPOEIVPLVQQQQPFGGQQQPF--PQQPFPQQFPFSPQPIYLQQLPF	61		
DB	18 PVP--QFPQPFSSQCP-----QQTFFQPDQTFHQDQQQCPQDQ--PQQQFLQPDQPF	66		
QY	62 P-QPWLPRYQPSFPPQQFPYQPC-----PQYSQQPQFISQQQAQ-----QQQQQQQQQQQ	110		
DB	67 PQQPQQQFPQ-----PQQPFPQTPQQQLFPQSQCPQQQSFQQQCFPQDPQPSFPPQ	122		
QY	111 QQQLIQQLLQQQLFCMNVLTQGNH-----IAHBSQVLQGSYVLTQLDELCCGHWQDPRQ	166		
DB	123 QPFLQPSLQQQVNCNKFLLQDQCKPVSIVSSVLSMIMPSQSCVQVQKQSCQQLAQIPQ	182		
QY	167 SQCCAIHNVVNAIILHQQKQQQQDPSQVSPQQPL-QQYPLQGSFPPSQNPQAGSVQ	225		
DB	183 LQCAIHVTHSHIIHQEQQQ-----GWHILLPYQQQGVQGGTL-----VQGGGIGITQ	230		
QY	226 PQQLQFEIRVLTQLPAMGNVTIAPCTI--APF-----GIFG 264			
DB	231 PQQPAQLAIRSLVLTQLPTMCNVNVPPKCSITKPFSSVAVGIGG 276			

RESULT 8  
US-10-474-955-99  
; Sequence 99, Application US/10474955  
; Publication No. US20040241161A1  
; GENERAL INFORMATION:



APPLICANT: Drifhout, Jan W.  
APPLICANT: Koning, Frits  
APPLICANT: McAdam, Stephan N.  
APPLICANT: Ludwig, Solid Magne  
TITLE OF INVENTION: METHODS AND MEANS FOR USE OF HLA-DQ RESTRICTED T-CELL RECEPTORS  
FILE REFERENCE: 2799/71244-PCT-US  
CURRENT APPLICATION NUMBER: US/10/474,955  
CURRENT FILING DATE: 2003-10-13  
NUMBER OF SEQ ID NOS: 137  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 99  
LENGTH: 279  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Amino acid sequence of GAMMA-4  
US-10-474-955-99

Query Match 36.3%; Score 521; DB 17; Length 279;  
Best Local Similarity 46.2%; Pred. No. 3.4e-33;  
Matches 132; Conservative 26; Mismatches 76; Indels 52; Gaps 14;  
DB 4 PVPQLQPNPSSQGPQEQVPLVQGGQPPGQQQQP--PQGPYPQPPFPSPQPYLQLQPF 61  
18 PVP--QPHQPSQGP-----QQTFFPQQTFFHQPQQQFPQ--PQQQFLQPPQPF 66  
QY 62 P-QPWLPPYPPQSPFPQPPYPPQ-----PQYSQPQPPISQQAQ-----QQQQQQQQQQ 110  
DB 67 PQGPQPPYPQ-----QPQPPFPQTQPPQQLFPQSQQPQQQFPQPPQPPQPPQPPQ 122  
QY 111 QQQIQQIQLQQQLPCMDVVLQGNH-----IAHRSQVLAQSTYQLQELCCQHLMOIPEQ 166  
DB 123 QPFFIQPSLQQQVNPCKNFLLQCKKPSVLSMSMWPQSDQVMRQSSCQQLAQIPQ 182  
QY 167 SQCCAIHNVNAIILHQQKQKQPPSSQVSFPQPL-QQYPLGGSPFRSQNPQAQGSVQ 225  
DB 183 LQCAIHHTVHSIIHQEQEQ-----GMHILLPLVQQQGVGGTL-----VGGGGIIG 230  
QY 226 PQQLPQFEIRNLALQTLPMKCNVYIAPYCTI--APF-----GIFG 264  
DB 231 PQQPAQLKAIKSLVLTPLPMCNVYVPECSIIKAPFSSVVAIGG 276

RESULT 9  
US-10-474-955-98  
Sequence 98, Application US/10474955  
Publication No. US20040241161A1  
GENERAL INFORMATION:  
APPLICANT: Drifhout, Jan W.  
APPLICANT: Koning, Frits  
APPLICANT: McAdam, Stephan N.  
APPLICANT: Ludwig, Solid Magne  
TITLE OF INVENTION: METHODS AND MEANS FOR USE OF HLA-DQ RESTRICTED T-CELL RECEPTORS  
FILE REFERENCE: 2799/71244-PCT-US  
CURRENT APPLICATION NUMBER: US/10/474,955  
CURRENT FILING DATE: 2003-10-13  
NUMBER OF SEQ ID NOS: 137  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 98  
LENGTH: 279  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Amino acid sequence of GAMMA-2  
US-10-474-955-98

Query Match 36.0%; Score 518; DB 17; Length 279;  
Best Local Similarity 45.8%; Pred. No. 5.8e-33;  
Matches 131; Conservative 26; Mismatches 77; Indels 52; Gaps 14;  
DB 4 PVPQLQPNPSSQGPQEQVPLVQGGQPPGQQQQP--PQGPYPQPPFPSPQPYLQLQPF 61

DB 18 PVP--QPHQPSQGP-----QQTFFPQQTFFHQPQQQFPQ--PQQQFLQPPQPF 66  
QY 62 P-QPWLPPYPPQSPFPQPPYPPQ-----PQYSQPQPPISQQAQ-----QQQQQQQQQQ 110  
DB 67 PQGPQPPYPQ-----QPQPPFPQTQPPQQLFPQSQQPQQQFPQPPQPPQPPQPPQ 122  
QY 111 QQQIQQIQLQQQLPCMDVVLQGNH-----IAHRSQVLAQSTYQLQELCCQHLMOIPEQ 166  
DB 123 QPFFIQPSLQQQVNPCKNFLLQCKKPSVLSMSMWPQSDQVMRQSSCQQLAQIPQ 182  
QY 167 SQCCAIHNVNAIILHQQKQKQPPSSQVSFPQPL-QQYPLGGSPFRSQNPQAQGSVQ 225  
DB 183 LQCAIHHTVHSIIHQEQEQ-----GMHILLPLVQQQGVGGTL-----VGGGGIIG 230  
QY 226 PQQLPQFEIRNLALQTLPMKCNVYIAPYCTI--APF-----GIFG 264  
DB 231 PQQPAQLKAIKSLVLTPLPMCNVYVPECSIIKAPFSSVVAIGG 276

RESULT 10  
US-10-474-955-100  
Sequence 100, Application US/10474955  
Publication No. US20040241161A1  
GENERAL INFORMATION:  
APPLICANT: Drifhout, Jan W.  
APPLICANT: Koning, Frits  
APPLICANT: McAdam, Stephan N.  
APPLICANT: Ludwig, Solid Magne  
TITLE OF INVENTION: METHODS AND MEANS FOR USE OF HLA-DQ RESTRICTED T-CELL RECEPTORS  
FILE REFERENCE: 2799/71244-PCT-US  
CURRENT APPLICATION NUMBER: US/10/474,955  
CURRENT FILING DATE: 2003-10-13  
NUMBER OF SEQ ID NOS: 137  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 100  
LENGTH: 279  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Amino acid sequence of GAMMA-3  
US-10-474-955-100

Query Match 36.0%; Score 518; DB 17; Length 279;  
Best Local Similarity 46.0%; Pred. No. 5.8e-33;  
Matches 131; Conservative 28; Mismatches 76; Indels 50; Gaps 14;  
DB 4 PVPQLQPNPSSQGPQEQVPLVQGGQPPGQQQQP--PQGPYPQPPFPSPQPYLQLQPF 61  
DB 18 PVP--QPHQPSQGP-----QQTFFPQQTFFHQPQQQFPQ--PQQQFLQPPQPF 66  
QY 62 P-QPWLPPYPPQSPFPQPPYPPQ-----PQYSQPQPPISQQAQ-----QQQQQQQQQQ 111  
DB 67 PQGPQPPYPQ-----QPQPPFPQTQPPQQLFPQSQQPQQQFPQPPQPPQPPQPPQ 123  
QY 112 QQQIQQIQLQQQLPCMDVVLQGNH-----IAHRSQVLAQSTYQLQELCCQHLMOIPEQ 167  
DB 124 PPFIQPSLQQQVNPCKNFLLQCKKPSVLSMSMWPQSDQVMRQSSCQQLAQIPQ 183  
QY 168 SQCCAIHNVNAIILHQQKQKQPPSSQVSFPQPL-QQYPLGGSPFRSQNPQAQGSVQ 226  
DB 184 LQCAIHHTVHSIIHQEQEQ-----GMHILLPLVQQQGVGGTL-----VGGGGIIG 231  
QY 227 PQQLPQFEIRNLALQTLPMKCNVYIAPYCTI--APF-----GIFG 264  
DB 232 QQPAQLKAIKSLVLTPLPMCNVYVPECSIIKAPFSSVVAIGG 276

RESULT 11  
US-10-739-930-9778  
Sequence 9778, Application US/10739930  
Publication No. US20040216190A1



```

GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
FILE REFERENCE: 38-21(53377)B
CURRENT APPLICATION NUMBER: US/10/739,930
CURRENT FILING DATE: 2003-12-18
NUMBER OF SEQ ID NOS: 11088
SEQ ID NO 9778
LENGTH: 304
TYPE: PRF
ORGANISM: Triticum aestivum
FEATURE:
OTHER INFORMATION: Clone ID: TRIAE-23APR03-C176_243.p
US-10-739-930-9778

Query Match
Best Local Similarity 42.1%; Score 445.5; DB 17; Length 304;
Matches 123; Conservative 34; Mismatches 76; Indels 57; Gaps 16;

QY 19 QGCVPLVQQQQQPFQGGQQQFPFPQGPYFPQGPFPFSSQGPYLQLQFPF-QEWLPYPQGSFPFPQ 77
DB 20 QMETSCISGLERPMQQLPLPPOQSFSSQGPFPFSQQQ--QPLPQGSFSSQGP-PPSQ 74
QY 78 QAPFQGPQYSGQGPISQQAQAGQQQQQ--QQQQQQQLLQI-----LQQQLIFC 126
DB 75 QPLSQGPFPSSQGGQGVLPFGQSPFSQQQLVLPQQQQQLVQQQIPIVQPSVQLQNP 134
QY 127 MDVVLQGH-NINH---ARQVLQOSTYQLIELCCGHLQIPDGSCQCAINHVHAI 180
DB 135 -KVFLLQCCQSPVAMPQRLARSGMQWQSSCHVWQQCCQLQIPEGSRYEKIALIYSII 193
QY 181 LHQGQK-----QQGQS-----SQVFEQGPQLQYPIFGSPFSQGNP 218
DB 194 LQEQGGQFVQPPQQQQLQSGGQGSQSSQSSQQLGQCSFQPPQQ--LQ--QPGQQQ 248
QY 219 Q--AGGS-VQPGQLPQFEERINLAIQTLFAMCNVYIAY--CTIAPFGIFGT 265
DB 249 QQVLGFPLQPIQIAHLEAVTISALRTPLTMCSSVNVPLYSATSVFEGV-GT 299

RESULT 12
US-10-739-930-9782
; Sequence 9782, Application US/10739930
; Publication No. US20040216190A1
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
FILE REFERENCE: 38-21(53377)B
CURRENT APPLICATION NUMBER: US/10/739,930
CURRENT FILING DATE: 2003-12-18
NUMBER OF SEQ ID NOS: 11088
SEQ ID NO 9782
LENGTH: 307
TYPE: PRF
ORGANISM: Triticum aestivum
FEATURE:
OTHER INFORMATION: Clone ID: TRIAE-23APR03-C176_269.p
US-10-739-930-9782

Query Match
Best Local Similarity 42.8%; Score 437.5; DB 17; Length 307;
Matches 125; Conservative 36; Mismatches 76; Indels 55; Gaps 16;

QY 13 PSCQGPQGPQVPLVQQQGF-----GQQQGFPPQGPYFPQGPFPFSSQGPYLQLQFPFGWLP 67
DB 27 PGLERPMQQLPLPPOQTFPQGPFLFSQQQ--QQLFPGQSPFSQGP-----PWCQQCP 78
QY 68 YPQGPFPFPQ-----YQPGQYSGPQGP-----SQQAQAGQQQQQQQQQQQI- 114
DB 79 FSCQGPPLPPOQFPSSQQLLVLPQ-QPFSQQQGVLPFPQSPFSFPQGGQQLVQQQIP 137

```

```

QY 111 -VQQLIIOOOLIPCMVVLVQCH--NIAH-----ASSQVLQSGSTYGLDELCCOHLMQIPBQS 167
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 138 VQPSILIQDLPNC-KYFLIQQCCSPVAMPQRLASQMLQSSCHVMQCCQQLPOLPQQS 196
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 168 QCAAIHNVHAITLHQQCKQKQOQPSQVSFOQPOLQYFLGQGSFRBSQNPQ----- 219
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 197 REVAIRAIYVITL--QEQCCVQGSISQOQOQPOQ---LQCVQSQPOQSOQQLGQOPOQ 251
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 220 ----AQGS-VQPOLPQPEIRNLALQTLPAKCNVYIAPY--CTIAFGIGFT 265
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 252 QQLAQSTFLQPHQIQLEVMVTSIRLILPMTGSNVVELYRTTTSVFGV-GT 302
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
US-10-739-930-9769
Sequence 9769, Application US/10739930
Publication No. US20040216190A1
GENERAL INFORMATION:
APPLICANT: Kowalic, David K.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
FILE REFERENCE: 38-21(53377)B
CURRENT APPLICATION NUMBER: US/10/739,930
CURRENT FILING DATE: 2003-12-18
NUMBER OF SEQ ID NOS: 11088
SEQ ID NO 9769
LENGTH: 244
TYPE: PRT
ORGANISM: Triticum aestivum
FEATURE:
OTHER INFORMATION: Clone ID: TRIAE-23APR03-C176_124.P
US-10-739-930-9769

Query Match 26.5%; Score 361; DB 17; Length 244;
Best Local Similarity 41.4%; Pred. No. 3,2e-22;
Matches 109; Conservative 33; Mismatches 61; Indels 60; Gaps 14;

QY 39 PQQPYR-QQP---PFSQQPYIQLQPFQPMPLPYFPQPSRFPPOQPYQPQP---QYQOPQ 92
DB 1 PQQPFPLQPSQSFMLQSQQPFVQ-----QPOQSPQPOQVQVQITSPAT 43
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 93 PISQQA-----QOQOQOQOQOQOQOQOQOQILQOQILQOQILPQMDVVLQCH--NIAH-- 138
DB 44 PTTIPAGKPTAPRPPQOQOQHQQLAQQQLFVQPSILQDLPNC-KYFLQOQCSPVAMPQ 102
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 139 --ARSOVLQOQSTYQLQELCCOHLMQIPBQSQCAIHNVAITLHQQCKQOQOQPSQVS 196
DB 103 RLRSQMLQOQSSCHVMQCCQCCOLPOLPQOBSRYAIRAIYVITL--QEQOQVQGSISQ 160
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 197 FQQLQYVFLGQGSFRBSQNPQ-----AQGS-VQPOLPQPEIRNLALQTLP 244
DB 161 QQQPQ---LQCVQSQPOQSOQQLGQOPQOQQLAQGFLQPHQIQLEVMVTSIALRILP 217
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 245 AMCNVYIAPY--CTIAFGIGFT 265
DB 218 TMCNVVPLYRTTTSVFGV-GT 239
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
US-10-425-115-200100
Sequence 200100, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kowalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Gao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326

```







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comphen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 17:02:00 ; Search time 15 seconds

(without alignments)  
1706.244 Million cell updates/sec

Title: US-10-089-700-3-w65

Perfect score: 1437

Sequence: 1 VAVVPQLQPNPNSQQQPOE.....CNVYIAPYCTIAPGIFGTN 266

Scoring table: BLOSUM62

Searched: 283416 seqs, 96216763 residues

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR79:\*

2: PIR1:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1361	94.7	286	2 S07923	alpha/beta-gliadin
2	1353	94.2	286	1 EEWTA	alpha/beta-gliadin
3	1309	91.1	288	2 T06282	alpha-gliadin prec
4	1264.5	88.0	291	2 T06498	alpha/beta-gliadin
5	1264.5	88.0	307	2 S10015	alpha/beta-gliadin
6	1260	87.7	296	2 S07361	alpha/beta-gliadin
7	1258	87.5	313	2 S07924	alpha/beta-gliadin
8	1254	87.3	296	2 A27319	gliadin - wheat
9	1245.5	86.7	319	2 A23364	alpha/beta-gliadin
10	1242.5	86.5	297	2 T02364	alpha/beta-gliadin
11	1232.5	85.8	326	2 T06500	alpha/beta-gliadin
12	1220	84.9	320	2 D22364	alpha/beta-gliadin
13	1203	83.7	320	2 E22364	alpha/beta-gliadin
14	1180	82.1	282	2 T06504	alpha/beta-gliadin
15	1155.5	80.4	292	2 B23364	alpha/beta-gliadin
16	586	40.8	327	2 JS0402	gamma-gliadin prec
17	546.5	38.0	291	1 EEWTA	gamma-gliadin B pr
18	538	37.4	302	2 JAO153	gamma-gliadin prec
19	490	34.1	305	2 S08312	gamma-hordein 1 pr
20	469	32.6	285	2 T05718	gamma-hordein 3 -
21	459.5	32.0	374	2 T05923	glutenin low molec
22	454.5	31.6	359	2 T06982	glutenin low molec
23	451.5	31.4	356	2 T05992	glutenin low molec
24	445.5	31.0	304	2 T06505	glutenin low molec
25	445	31.0	264	2 S07505	B3-hordein (clone
26	444	30.9	293	2 S07365	hordein B1 - barley
27	442	30.8	285	2 S20853	glutenin low molec
28	441.5	30.7	298	2 T06980	glutenin low molec
29	440	30.6	286	2 T05910	glutenin low molec

30	438	30.5	303	2 T06981	low-molecular-weight
31	437.5	30.4	290	2 S20519	hordein B precursor
32	437.5	30.4	307	2 S04325	glutenin low molec
33	425	29.6	251	2 PS0094	gamma-gliadin prec
34	423.5	29.5	221	2 T04474	B1 hordein - barley
35	414	28.8	295	1 JN0696	glutenin low molec
36	410.5	28.6	209	2 B36433	avenin precursor
37	407	28.3	220	2 JQ1046	avenin precursor (
38	388.5	27.0	194	2 A23277	gamma-secalin - ry
39	383	26.7	243	2 S07976	B1-hordein (clone
40	382.5	26.6	222	2 JQ1047	avenin precursor (
41	381	26.5	244	2 S07398	gamma-gliadin B pr
42	376	26.2	182	2 JG0015	avenin N9 - oat
43	374.5	26.1	323	2 T06506	glutenin low molec
44	374	26.0	261	2 S57655	glutenin low molec
45	369.5	25.7	276	2 S57656	glutenin low molec

## ALIGNMENTS

## RESULT 1

S07923

alpha/beta-gliadin precursor - wheat

C/Species: Triticum aestivum (common wheat)

C/Date: 08-Jun-1994 #sequence\_revision 01-Dec-1995 #text\_change 20-Aug-1999

C/Accession: S07923

R/Submitter: Smith, M.; Rafalski, J.A.; Sugiyama, T.; Stoll, M.; Soell, D.

Nucleic Acids Res. 13, 3905-3916, 1985

A/Title: Conservation and variability of wheat alpha/beta-gliadin genes.

A/Reference number: S07361; MUID:85242077; PMID:3839304

A/Accession: S07923

A/Status: preliminary; translation not shown

A/Molecule type: DNA

A/Residues: 1-286 <SUM>

A/Cross-references: EMBL:X02539; NID:g21760; PIDN:CAA26384.1; PID:g21761

C/Superfamily: gliadin

Query Match	Score	1361;	DB 2;	Length	286;
Best Local Similarity	95.98;	Pred. No.	1.8e-84;		
Matches	255;	Conservative	0;	Mismatches	11;
		Indices	0;	Gaps	0;
QY	1	VAVVPQLQPNPNSQQQPOEQLVVOQQCPFGQQQFPPOQPYPOQPFPSQPYLOLP	60		
DB	21	VAVVPQLQPNPNSQQQPOEQLVVOQQCPFGQQQFPPOQPYPOQPFPSQPYLOLP	80		
QY	61	PPQFMFPPOQSPPOQPYPOQPOYSQPOPISSQQAQQQQQQQQQQQIIOQIIQ	120		
DB	81	FPQFMFPPOQSPPOQPYPOQPOYSQPOPISSQQAQQQQQQQQQQQIIOQIIQ	140		
QY	121	QQLIPCMQVVLQOHNIHAHRSQVLOQSTYQLQELCOHMOIPESQCOAHINVAHII	180		
DB	141	QQLIPCMQVVLQOHNIHAHRSQVLOQSTYQLQELCOHMOIPESQCOAHINVAHII	200		
QY	181	LHQQKXQQQSSQVSPQPLQOYPLQGSFRSQNPQAGSVPOQLPQFEIRNLAL	240		
DB	201	LHQQKXQQQSSQVSPQPLQOYPLQGSFRSQNPQAGSVPOQLPQFEIRNLAL	260		
QY	241	QTLPMQVYIAPYCTIAPGIFGTN	266		
DB	261	QTLPMQVYIAPYCTIAPGIFGTN	286		

## RESULT 2

EEWTA

alpha/beta-gliadin precursor - wheat

N/Alternate names: prolamin

C/Species: Triticum aestivum (common wheat)

C/Date: 28-May-1986 #sequence\_revision 28-May-1986 #text\_change 09-Jul-2004

C/Accession: A03354

R/Rafalski, J.A.; Schaefer, K.; Metzler, M.; Peterson, D.M.; Hedgcock, C.; Stoll, D.G.

EMBO J. 3, 1409-1415, 1984

A/Title: Developmentally regulated plant genes: the nucleotide sequence of a wheat glia



141. LQQQLTFCMEVLLQQHNIARGRSOVLQOSTVOLQBLCCGHLWQIPEKLQCGA1HNVVA 200

Query Match	88.0%; Score 1264.5; DB 2; Length 307;
<p> <i>Fl213011</i> Product: <i>alpha/beta-galactin</i> #baculus predicted size: 307aa            100% identical to: <i>Fl213011</i> </p>	

1998







QY 112 QQLIQQQLIPCMQDVVLCQHNIHARSQVLAQSTYQLLQELCCQHLMOIEGSCCA 171  
 DB 141 QQLIPQLIQQQLIPCMQDVVLCQHNIHARSQVLAQSTYQLLQELCCQHLMOIEGSCCA 200  
 QY 172 IHHVVAHIIHQQQKQQQSSQVSPQPIQQYPLGGSGRPSQONPQAQGSVQPOLPQ 231  
 DB 201 IHHVVAHIIHQQQKQQQSSQVSPQPIQQYPLGGSGRPSQONPQAQGSVQPOLPQ 259  
 QY 232 FEIRNLALQTLPMQCNVYIAPYC--TIAPGIFGTN 266  
 DB 260 FEIRNLALQTLPMQCNVYIAPYC--TIAPGIFGTN 296

## RESULT 9

A22364  
 alpha/beta-gliadin precursor (clone A42) - wheat  
 Name alternate names: prolamin  
 C:Species: Triticum aestivum (common wheat)  
 C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 09-Jul-2004  
 G:Accession: A22364  
 R:Okita, T.W.; Cheesbrough, V.; Reeves, C.D.  
 J: Biol. Chem. 260, 8203-8213, 1985  
 A:Title: Evolution and heterogeneity of the alpha/beta-type and gamma-type gliadin DNA  
 A:Reference number: A92541; MUID:85234522; PMID:2989281  
 A:Accession: A22364  
 A:Molecule type: mRNA  
 A:Residues: 1-319 <OK>  
 A:Cross-references: UNIPROT:P04725; GB:M11073; NID:g170715; PIDN:AAA34278.1; PID:g170716  
 G:Superfamily: gliadin  
 F:21-319/Product: signal sequence #status predicted <SIG>  
 F:21-319/Product: alpha/beta-gliadin #status predicted <MAT>

Query Match 86.7%; Score 1245.5; DB 2; Length 319;  
 Best Local Similarity 81.9%; Pred. No. 1.1e-76;  
 Matches 245; Conservative 5; Mismatches 16; Indels 33; Gaps 4;  
 QY 1 VRVVPVQLQPNPNSQQQPOEQVPLVQQQOFPGQQQPPPOQPYPOPFPSQQPYLQLQP 60  
 DB 21 VRVVPVQLQPNPNSQQQPOEQVPLVQQQOFPGQQQPPPOQPYPOPFPSQQPYLQLQP 80  
 QY 61 FPQ-----PWLPYPOQSPFPPOQPYPOPOPSQPPQPISSQQQA-QQQQQQQQQQQQQQ 114  
 DB 81 FPQPPFPPLPYPOQSPFPPOQPYPOPOPSQPPQPISSQQQA-QQQQQQQQQQQQQQ 140  
 QY 115 LQQLIQQQLIPCMQDVVLCQHNIHARSQVLAQSTYQLLQELCCQHLMOIEGSCCA 174  
 DB 141 LQQLIQQQLIPCMQDVVLCQHNIHARSQVLAQSTYQLLQELCCQHLMOIEGSCCA 200  
 QY 175 VVAHIIH-----QQQKQQQSSQVSPQPIQQYPLGGSGRPSQONPQAQGSVQ 209  
 DB 201 VVAHIIHQQQKQQQSSQVSPQPIQQYPLGGSGRPSQONPQAQGSVQ 260  
 QY 210 SFRPSQONPQAQGSVQPOLPQFEIRNLALQTLPMQCNVYIAPYC--TIAPGIFGTN 266  
 DB 261 SFRPSQONPQAQGSVQPOLPQFEIRNLALQTLPMQCNVYIAPYC--TIAPGIFGTN 319

## RESULT 10

A22364  
 alpha/beta-gliadin precursor (clone A212) - wheat  
 C:Species: Triticum aestivum (common wheat)  
 C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 09-Jul-2004  
 G:Accession: A22364  
 R:Okita, T.W.; Cheesbrough, V.; Reeves, C.D.  
 J: Biol. Chem. 260, 8203-8213, 1985  
 A:Title: Evolution and heterogeneity of the alpha/beta-type and gamma-type gliadin DNA  
 A:Reference number: A92541; MUID:85234522; PMID:2989281  
 A:Accession: A22364  
 A:Molecule type: mRNA  
 A:Residues: 1-319 <OK>  
 A:Cross-references: UNIPROT:P04722  
 G:Superfamily: gliadin  
 F:1-20/Domain: signal sequence #status predicted <SIG>

F:21-319/Product: alpha/beta-gliadin #status predicted <MAT>  
 Query Match 86.5%; Score 1242.5; DB 2; Length 319;  
 Best Local Similarity 80.9%; Pred. No. 1.8e-76;  
 Matches 242; Conservative 7; Mismatches 17; Indels 33; Gaps 4;

QY 1 VRVVPVQLQPNPNSQQQPOEQVPLVQQQOFPGQQQPPPOQPYPOPFPSQQPYLQLQP 60  
 DB 21 VRVVPVQLQPNPNSQQQPOEQVPLVQQQOFPGQQQPPPOQPYPOPFPSQQPYLQLQP 80  
 QY 61 FPQ-----PWLPYPOQSPFPPOQPYPOPOPSQPPQPISSQQQA-QQQQQQQQQQQQQQ 114  
 DB 81 FPQPPFPPLPYPOQSPFPPOQPYPOPOPSQPPQPISSQQQA-QQQQQQQQQQQQQQ 140  
 QY 115 LQQLIQQQLIPCMQDVVLCQHNIHARSQVLAQSTYQLLQELCCQHLMOIEGSCCA 174  
 DB 141 LQQLIQQQLIPCMQDVVLCQHNIHARSQVLAQSTYQLLQELCCQHLMOIEGSCCA 200  
 QY 175 VVAHIIH-----QQQKQQQSSQVSPQPIQQYPLGGSGRPSQONPQAQGSVQ 209  
 DB 201 VVAHIIHQQQKQQQSSQVSPQPIQQYPLGGSGRPSQONPQAQGSVQ 260  
 QY 210 SFRPSQONPQAQGSVQPOLPQFEIRNLALQTLPMQCNVYIAPYC--TIAPGIFGTN 266  
 DB 261 SFRPSQONPQAQGSVQPOLPQFEIRNLALQTLPMQCNVYIAPYC--TIAPGIFGTN 319

## RESULT 11

T06500  
 alpha/beta-gliadin A-IV precursor - wheat  
 C:Species: Triticum aestivum (common wheat)  
 C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
 G:Accession: T06500  
 R:Okita, T.W.; Cheesbrough, V.; Reeves, C.D.  
 J: Biol. Chem. 260, 8203-8213, 1985  
 A:Title: Evolution and heterogeneity of the alpha/beta-type and gamma-type gliadin DNA  
 A:Reference number: A92541; MUID:85234522; PMID:2989281  
 A:Accession: T06500  
 A:Molecule type: mRNA  
 A:Residues: 1-297 <OK>  
 A:Cross-references: UNIPROT:P04724; EMBL:M11075; NID:g170723; PIDN:AAA34282.1; PID:g170724  
 G:Superfamily: gliadin  
 C:Keywords: seed; storage protein  
 F:1-20/Domain: signal sequence #status predicted <SIG>  
 F:1-297/Product: alpha/beta-gliadin A-IV #status predicted <MAT>

Query Match 85.8%; Score 1232.5; DB 2; Length 297;  
 Best Local Similarity 84.3%; Pred. No. 7.7e-76;  
 Matches 236; Conservative 9; Mismatches 18; Indels 17; Gaps 3;

QY 1 VRVVPVQLQPNPNSQQQPOEQVPLVQQQOFPGQQQPPPOQPYPOPFPSQQPYLQLQP 60  
 DB 21 VRVVPVQLQPNPNSQQQPOEQVPLVQQQOFPGQQQPPPOQPYPOPFPSQQPYLQLQP 80  
 QY 61 F-----PQWLPYPOQSPFPPOQPYPOPOPSQPPQPISSQQQA-QQQQQQQQQQQQQQ 113  
 DB 81 FPQPPFPPLPYPOQSPFPPOQPYPOPOPSQPPQPISSQQQA-QQQQQQQQQQQQQQ 137  
 QY 114 IQQQLIQQQLIPCMQDVVLCQHNIHARSQVLAQSTYQLLQELCCQHLMOIEGSCCA 173  
 DB 138 IQQQLIQQQLIPCMQDVVLCQHNIHARSQVLAQSTYQLLQELCCQHLMOIEGSCCA 197  
 QY 174 NVVAHIIH-----QQQKQQQSSQVSPQPIQQYPLGGSGRPSQONPQAQGSVQ 226  
 DB 198 NVVAHIIHQQQKQQQSSQVSPQPIQQYPLGGSGRPSQONPQAQGSVQ 257  
 QY 227 QQLPQFEIRNLALQTLPMQCNVYIAPYC--TIAPGIFGTN 266  
 DB 258 QQLPQFEIRNLALQTLPMQCNVYIAPYC--TIAPGIFGTN 297

## RESULT 12







A:Residues: 1-292 <OKI>  
A:Cross-references: UNIPROT:P04721  
C:Superfamily: gliadin  
F:1-20/Domain: signal sequence #status predicted <SIG>  
F:21-292/Product: alpha/beta-gliadin #status predicted <MAT>

Query Match 80.4%; Score 1155.5; DB 2; Length 292;  
Best Local Similarity 84.1%; Pred. No. 1,1e-70;  
Matches 227; Conservative 3; Mismatches 9; Indels 31; Gaps 4;

```
QY 1 VRVVPQLQPNPSCQCPQECVPLVQCCQFPQCCQFPQCPYQPPQPPSQQPYLQLCP 60
DB 21 VRVVPQLQLNPSQCCQPEQVPLVQCCQFLGQCCPPFPQPPYQPPPSQQPYLQLCP 80
QY 61 F--PQW---LPYQPSFPQCPYQPPQYSPQPPIS-QQQACQCCQCCQCCQCCQCC 114
DB 81 FLQPPFPQPLPSPQCPFRFPQPPQPPQPPQPPQPPISQCCQCCQCCQCCQCCQCC 140
QY 115 LQCLLQQLLPCMDVVLQGHVIAHRSQVLQOSTYLLDELCCQHLWQIPESQCCAIHN 174
DB 141 LQCLLQQLLPCMDVVLQGHVIAHRSQVLQOSTYLLDELCCQHLWQIPESQCCAIHN 200
QY 175 VVHAIIILH-----QQHQQQQPSSQVSFQQPLQYPLGG 209
DB 201 VVHAIIILHQQQQQQEQKQLQCCQCCQCCQCCQCCQCCQCCQCCQCCQCCQCCQCC 260
QY 210 SFRPQQNPQAQGSVQPPQLPQEEIRNLA 239
DB 261 SFRPQQNPQAQGSVQPPQLPQEEIRNLA 290
```

Search completed: December 14, 2004, 17:27:00  
Job time : 16 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 17:01:05 ; Search time 79 Seconds  
(without adjustments)

Title: US-10-089-700-3-W65

Sequence: 1 VRVFPQLQPQNPSQQQPE.....CNVYIAPYCTIAPFGIFGTN 266

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Minimum Match 0%

```
Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1365	95.0	269	2	Q9M417	Q9m417 triticum ae
2	1353	94.2	286	1	GD40_WHEAT	P02963 triticum ae
3	1353	94.2	286	2	AA965525	AA965525 triticum ae
4	1351	94.0	288	2	Q9ZP09	Q9zP09 triticum ae
5	1347.5	93.8	274	2	Q9M4M5	Q9m4M5 triticum ae
6	1347.5	93.8	276	2	Q9M4M2	Q9m4M2 triticum ae
7	1347.5	93.4	277	2	Q41509	Q41509 triticum ae
8	1341	93.3	277	2	Q9M4I8	Q9m4I8 triticum ae
9	1338.5	93.1	276	2	Q9M4M0	Q9m4M0 triticum ae
10	1338.5	92.9	289	2	Q41531	Q41531 triticum ae
11	1327.5	92.4	270	2	Q9M4I9	Q9m4I9 triticum ae
12	1326.5	92.3	278	2	Q9M4I1	Q9m4I1 triticum ae
13	1312.5	91.3	259	2	Q41533	Q41533 triticum ae
14	1309	91.1	288	2	Q41530	Q41530 triticum ae
15	1296.5	90.2	287	2	Q41528	Q41528 triticum ae
16	1264.5	88.0	290	2	Q9M4I6	Q9m4I6 triticum ae
17	1264.5	88.0	291	1	GD42_WHEAT	P04722 triticum ae
18	1264.5	88.0	307	1	GD46_WHEAT	P48573 triticum ae
19	1260	87.7	296	1	GD47_WHEAT	P04726 triticum ae
20	1259	87.6	313	1	GD47_WHEAT	P04727 triticum ae
21	1258	87.5	313	2	Q41546	Q41546 triticum ae
22	1254	87.3	296	2	Q41632	Q41632 triticum ae
23	1251	87.1	273	2	Q9M4M4	Q9m4M4 triticum ae
24	1246	86.7	318	2	Q41545	Q41545 triticum ae
25	1245.5	86.7	319	1	GD45_WHEAT	P4725 triticum ae
26	1243	86.5	313	2	Q41529	Q41529 triticum ae
27	1232.5	86.8	297	1	GD44_WHEAT	P04724 triticum ae
28	1197.5	83.3	262	1	GD41_WHEAT	P04731 triticum ae
29	1194	83.1	265	2	Q9M4M3	Q9m4M3 triticum ae
30	1190	82.8	265	2	Q9M4I6	Q9m4I6 triticum ae
31	1180	82.1	282	1	GD43_WHEAT	P04723 triticum ae

32	933	64.9	186	1	GD8A	WHEAT	P04428	tritium ae
33	591.5	41.2	455	2	Q9PFA1		Q9PFA1	secal cere
34	587	40.8	308	2	Q9M6P7	tritium ae	Q9M6P7	tritium ae
35	587	40.8	311	2	Q9E6W9	tritium ae	Q9E6W9	tritium ae
36	587	40.8	327	2	Q94G91	tritium ae	Q94G91	tritium ae
37	586	40.8	327	1	GD8B	WHEAT	P08453	tritium ae
38	578	40.2	337	2	Q94G96	tritium ae	Q94G96	tritium ae
39	566	39.4	300	2	Q9EFA8	tritium ae	Q9EFA8	aeq11ops
40	566	39.4	300	2	Q9FVA1	tritium ae	Q9FVA1	aeq11ops
41	555.5	38.7	274	2	Q6ERX0	tritium ae	Q6ERX0	tritium ae
42	555.5	38.7	282	2	Q6ESW7	tritium ae	Q6ESW7	tritium ae
43	553.5	38.7	288	2	Q94G92	tritium ae	Q94G92	tritium ae
44	553.5	38.5	275	2	Q6ERW4	tritium ae	Q6ERW4	tritium ae
45	552	38.4	275	2	Q6ERW2	tritium cu	Q6ERW2	tritium cu

## ALIGNMENTS

ID	Q9M4L7	PRELIMINARY;	PRT,	269 AA.
AC	Q9M4L7			
DT	01-OCT-2000	(TREMBREL. 15)	Created	
DT	01-OCT-2000	(TREMBREL. 15)	Last sequence update)	
DT	01-MAR-2004	(TREMBREL. 26)	Last annotation update)	
DB	Alpha-gliadin.			
OS	Triticum aestivum (Wheat).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
CC	Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;			
NCBI_TaxID=4565;	Triticeae; Triticum.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Mjoelner; TISSUE=Endosperm;			
RA	Arentz-Hansen E.H., Mcadam S.N., Molberg O., Kristiansen C.,			
RL	Solid LM.			
DR	Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.			
DR	EMBL; AJ133611; CAB76963.1; -			
DR	GO; GO:0045735; F:nutrient reservoir activity; IEA.			
DR	InterPro; IPR003612; AAI.			
DR	InterPro; IPR001376; Gliadin.			
DR	InterPro; IPR001954; Glia-glutenin.			
DR	Pfam; PF00234; TRYP_alpha_amy1; 1.			
DR	PRINTS; PR00208; GLIADGLUTEN.			
DR	PRINTS; PR00209; GLUTADIN.			
DR	SMART; SM00499; AAI; 1.			
FT	CHAIN			
SQ	SEQUENCE	269 AA;	269	alpha-gliadin.
		1	31292 MW;	87127D6FD15EC78B CRC64;
	Query Match	95.0%;	Score 1365;	DB 2:
	Best Local Similarity	96.2%;	Fred. No.3.5e-76;	
	Matches	256;	Conservative	0;
			Mismatches	10;
			Indels	0;
			Gaps	0
QY	1 VRVFPQQPQPQNPSQCQPEQEVPVLVQQCQPFQGCQFPPQCFYPQOPDFSPSGPYQLQDP			60
Db	2 VRVLPQIQPNQNSQCCEVEVPLVQQCQFLGCGCQFFPFCQFPYPAQPFSSQLPVLIQDP			61
QY	61 FPQPMPLPFGQGSPFPQCPYFPQPGYSQPOGPISQCAAQGCCGGCCGQHIIHQILQ			120
Db	62 FPQPLPFSQPFPPQCPYPQPQPSQPCQPISQCCCCQQCCGQGHIIHQILQ			121
QY	121 QCLFPCMVMVLAQHNIHARSQVLQGSFYQLQELCCGHLWCIPESQSCQAHHNVHAII			180
Db	122 QCLIPCMMVMVAQHNIHARGSVQLQGSFYQLQELCCGHLWCIPESQSCQAHHNVHAII			181
QY	181 LHQQCKKQQQPSQSVCVFQGPLQQYFLGGSTRFPSQNNQAQGVSVPOOLPFEETRNAL			240
Db	182 LHQQCKKKQQQSSQVSFOQPLQQXYFLGGSTRFPSQNNQAQGVSVPOOLPFEETRNLAL			241
QY	241 QTLPAMCMVTATPYCTIAAPFGIFGTN			266
Db	242 QTLPAMCMVTITPTCTIAAPFGIFGTN			267



RESULT 2  
 GDAD WHEAT STANDARD; PRT; 286 AA.  
 AC P02863;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Alpha/beta-gliadin precursor (Prolamin).  
 OS Triticum aestivum (wheat).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
 OC Triticeae; Triticum.  
 GX NCBI\_TaxID=4565;  
 RX F1  
 RE SEQUENCE FROM N.A.  
 RC STRAIN=cv. Newton;  
 RX MBLIN=84261434; PubMed=6204862;  
 RA Rafalski J.A., Scheets K., Metzler M., Peterson D.M., Hedgcock C.,  
 RA Sell D.G.;  
 RT "Developmentally regulated plant genes: the nucleotide sequence of a  
 RT wheat gliadin genomic clone.";  
 RL EMBO J. 3:1409-1415(1984).  
 RX F2  
 RE SEQUENCE FROM N.A. (CLONE PM8233).  
 RX MBLIN=85242077; PubMed=3839304;  
 RA Summer-Smith M., Rafalski J.A., Sugiyama T., Stoll M., Soell D.;  
 RT "Conservation and variability of wheat alpha/beta-gliadin genes.";  
 RL Nucleic Acids Res. 13:3905-3916(1985).  
 RX F3  
 RE SEQUENCE FROM N.A.  
 RX MBLIN=85062803; PubMed=6095191;  
 RA Anderson O.D., Lits J.C., Gauthier M.F., Greene F.C.;  
 RT "Nucleic acid sequence and chromosome assignment of a wheat storage  
 RT protein gene.";  
 RL Nucleic Acids Res. 12:8129-8144(1984).  
 CC -1- FUNCTION: Gliadin is the major seed storage protein in wheat.  
 CC -1- PMR: Substrate of transglutaminase (By similarity).  
 CC -1- ALLERGEN: Causes an allergic reaction in human. Is the cause of  
 CC the celiac disease, also known as celiac sprue or gluten-sensitive  
 CC enteropathy (By similarity).  
 CC -1- MISCELLANEOUS: The alpha/beta-gliadins can be divided into 5  
 CC homology classes. Sequence divergence between the classes is due  
 CC to single base substitutions and to duplications or deletions  
 CC within or near direct repeats. There are more than a 100 copies of  
 CC the gene for alpha/beta-gliadin per haploid genome.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL: X00627; CAA5261.1; -  
 CC EMBL: X03076; AAA34280.1; -  
 CC EMBL: X02539; CAA26384.1; -  
 CC EMBL: X01130; CAA25593.1; -  
 CC PIR: A03354; EEMTA.  
 CC InterPro: IPR003612; AAI.  
 CC InterPro: IPR001376; Gliadin.  
 CC InterPro: IPR001954; Gliadin.  
 CC Pfam: PF00234; Tryptophan alpha-amylase 1.  
 CC PRINTS: PR00208; GLIADGUTEN.  
 CC PRINTS: PR00209; GLIADIN.  
 CC SMART: SM00499; AAI; 1.  
 CC Allergen: Multigene family; Repeat; Seed storage protein; Signal.  
 CC SIGNAL  
 CC CHAIN 21 286 Alpha/beta-gliadin.  
 CC VARIANT 37 37 L -> Q (in clone PM8233 and in Ref. 3).  
 CC VARIANT 93 93 P -> Q (in clone PM8233).

FT VARIANT 193 194 HN -> LK (in Ref. 3).  
 SQ SEQUENCE 286 AA; 32949 MW; ESECFABBE29E10C6 CRC64;  
 Query Match 94.2%; Score 1353; DB 1; Length 286;  
 Best Local Similarity 95.5%; Pred. No. 2e-75;  
 Matches 254; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
 QY 1 VRVVPOLQPNPSSQGFQEQVPLVQQQDFEGQQQFPFQQPYPQPPFSSQPPYLQLP 60  
 DB 21 VRFVPQLQPNPSSQGLFQEQVPLVQQQDFEGQQQFPFQQPYPQPPFSSQPPYLQLP 80  
 QY 61 FPPQWLPPYPQPSPPQPPYPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 120  
 DB 81 FPPQWLPPYPQPSPPQPPYPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 140  
 QY 121 QQLIPCMQVVLQOHNIARSGQVLAQSTYQLLQELCCQHLWQIPESQCAIHNVYAAII 180  
 DB 141 QQLIPCMQVVLQOHNIARSGQVLAQSTYQLLQELCCQHLWQIPESQCAIHNVYAAII 200  
 QY 181 LHQQKQKQKQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSS 240  
 DB 201 LHQQKQKQKQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSS 260  
 QY 241 QTLPMCNVYIAPYCTIAPFGIFGTN 266  
 DB 261 QTLPMCNVYIAPYCTIAPFGIFGTN 286  
 RESULT 3  
 AAA96525  
 ID AAA96525 PRELIMINARY; PRT; 286 AA.  
 AC AAA96525;  
 DT 02-MAR-2004 (TRENDEL 27, Created)  
 DT 02-MAR-2004 (TRENDEL 27, Last sequence update)  
 DT 02-MAR-2004 (TRENDEL 27, Last annotation update)  
 DE Alpha-gliadin storage protein.  
 OS Triticum aestivum (wheat).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Triticeae.  
 RX NCBI\_TaxID=4565;  
 RN (1)  
 RE SEQUENCE FROM N.A.  
 RC STRAIN=Cheyenne;  
 RA Anderson O.D.;  
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U51307; AAA96525.1; -  
 SQ SEQUENCE 286 AA; 32949 MW; ESECFABBE29E10C6 CRC64;  
 Query Match 94.2%; Score 1353; DB 2; Length 286;  
 Best Local Similarity 95.5%; Pred. No. 2e-75;  
 Matches 254; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
 QY 1 VRVVPOLQPNPSSQGFQEQVPLVQQQDFEGQQQFPFQQPYPQPPFSSQPPYLQLP 60  
 DB 21 VRFVPQLQPNPSSQGLFQEQVPLVQQQDFEGQQQFPFQQPYPQPPFSSQPPYLQLP 80  
 QY 61 FPPQWLPPYPQPSPPQPPYPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 120  
 DB 81 FPPQWLPPYPQPSPPQPPYPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 140  
 QY 121 QQLIPCMQVVLQOHNIARSGQVLAQSTYQLLQELCCQHLWQIPESQCAIHNVYAAII 180  
 DB 141 QQLIPCMQVVLQOHNIARSGQVLAQSTYQLLQELCCQHLWQIPESQCAIHNVYAAII 200  
 QY 181 LHQQKQKQKQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSS 240  
 DB 201 LHQQKQKQKQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSSQSS 260  
 QY 241 QTLPMCNVYIAPYCTIAPFGIFGTN 266  
 DB 261 QTLPMCNVYIAPYCTIAPFGIFGTN 286











CC Triticum aestivum (Wheat).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
 CC Triticaceae; Triticum.  
 OC NCBI\_TaxID=4555;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Mjoeiner; TISSUE=Endosperm;  
 RA Arentz-Hansen E.H., McAdam S.N., Molberg O., Kristiansen C.,  
 RA Solid L.N.,  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ133608; CAB76960.1; -  
 DR GO; GO:0045735; F:nutrient reservoir activity; IEA.  
 DR InterPro; IPR003612; AAI.  
 DR InterPro; IPR001376; Gliadin.  
 DR InterPro; IPR001954; Glia\_glutenin.  
 DR Pfam; PF00234; TRY\_alpha\_amyl\_1.  
 DR PRINTS; PR00208; GLIADIN.  
 DR PRINTS; PR00209; GLIADIN.  
 DR SMART; SM00499; AAI; 1.  
 DR CHAIN 1 276 alpha-gliadin.  
 FT SEQUENCE 276 AA; 32211 MW; 6A2E9723D42B100A CRC64;  
 SQ  
 Query Match 93.1%; Score 1338.5; DB 2; Length 276;  
 Best Local Similarity 92.7%; Pred.No.1.5e-74;  
 Matches 253; Conservative 2; MisMatches 11; Indels 7; Gaps 1  
 QY 1 VAVVPQLOLPQNPSSQCOQPOEQVPLVYQQQCFPPQCOQCOFPFPQCFYPQPOFPFPSSQCPYLQLOLP 60  
 DB 2 VAVVPQLOLPQNPSSQCOQPOEQVPLVYQQQCFPPQCOQCFPPQCFYPQPOFPFPSSQCPYLQLOLP 61  
 QY 61 FFPQWLPYPPQPSFFPPQCFYPQPOFPQYSCPPQCFISQQAQOQQOQQOQQOQ-----QQ 113  
 DB 62 FFPQPLSYSPQPOFFPPQCFYPQPOFPQYSCPPQCFISQQAQOQQOQQOQQOQ-----QQ 121  
 QY 114 IIQQIIQQOQLICMDVYLQGNHNHARSGVLTQOSTYQLLQELCCCHLMQIPESQCCQAIH 173  
 DB 122 IIQQIIQQOQLICMDVYLQGNHNHARSGVLTQOSTYQLLQELCCCHLMQIPESQCCQAIH 181  
 QY 174 NVVHAIIILHQQKQKQOQPSQSVSFQOPLQOYPLQGSGFFPSSQGNPQAQGSVQPOQLQFQE 233  
 DB 182 NVVHAIIILHQQKQKQOQPSQSVSFQOPLQOYPLQGSGFFPSSQGNPQAQGSVQPOQLQFQE 241  
 QY 234 EIRNMLAQTLPMQCNVYIAPYCTIAPFGIFGTN 266  
 DB 242 EIRNMLAQTLPMQCNVYIAPYCTIAPFGIFGTN 274  
 QY  
 DB  
 RESULT 10  
 Q41531  
 ID Q41531 PRELIMINARY; PRT; 289 AA.  
 AC Q41531;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-VAR-2004 (TREMBlrel. 26, Last annotation update)  
 OS Alpha-gliadin storage protein.  
 OS Triticum aestivum (Wheat).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
 OC Triticaceae; Triticum.  
 OC NCBI\_TaxID=4555;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Cheyenne;  
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U51306; AAA96524.1; -  
 DR PIR; S13333; S13333.  
 DR GO; GO:0045735; F:nutrient reservoir activity; IEA.  
 DR InterPro; IPR003612; AAI.  
 DR InterPro; IPR001376; Gliadin.  
 DR InterPro; IPR001954; Glia\_glutenin.  
 DR Pfam; PF00234; TRY\_alpha\_amyl\_1.  
 DR

```

DR PRINTS; PR00208; GLIADGLUTEN.
DR PRINTS; PR00209; GLIADIN.
DR SMART; SMO0499; AAI: 1.
SQ SEQUENCE 289 AA; 33349 MW; 5F577C9CD63874FA CRC64;

Query Match
Best Local Similarity 92.9%; Score 1334.5; DB 2; Length 289;
Matches 253; Conservative 1; Mismatches 12; Indels 3; Gaps 1.

QY 1 VRVVPQLOPQNPBQQQPOEQQVPLVQQQQQFFGQQQQQFPPOQYPQOPQPFPSQQPYLQLOP 60
DB 21 VRVSPQLOPQNPBQQQPOEQQVPLVQQQQQFFGQQQQFPFPQYPQLOQPFPSQQPYLQLOP 80
QY 61 FPOEWLPPYQPSQSPFPQYPQOPQYSPQOPIS--QQQAQQQQQQQQQQQQQ11QQ 117
DB 81 FPGQLPFSQQPFRFPQYPQOPQYSPQOPISQQQQQQQQQQQQQQQQQQQQ11QQ 140
QY 118 ILQQQLIFCMQVYLCQHNIARBSQVLAQSTYYQLLDELCOHLMQIPBQSCQALHNVAH 177
DB 141 ILQQQLIFCMQVYLCQHNIARBSQVLAQSTYYQLLRELCCQHLMQIPBQSCQALHNVAH 200
QY 178 AIIHQQKQQQSQSSQVSPQOPQYPLQGGSFRRPSQNPQAQSVQPOQLPQEEIRN 237
DB 201 AIIHQQKQQQSQSSQVSPQOPQYPLQGGSFRRPSQNPQAQSVQPOQLPQEEIRN 260
QY 238 IALQTLPMQNVYIAPYCTIAPFGIFGN 266
DB 261 IALQTLPMQNVYIAPYCTIAPFGIFGN 289

RESULT 11
Q9M4I9 PRELIMINARY; PRT; 270 AA.
ID Q9M4I9
AC Q9M4I9;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Alpha-gliadin.
OS Triticum aestivum (wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
OC Triticeae; Triticum.
OX NCBI_Taxid=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hjoesner; TISSUE=Endosperm;
RC Arentz-Hansen E.H.; McAdam S.N.; Wolberg O.; Kristiansen C.;
RL Solid L.M.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ133609; CAB76961.1; -.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR003612; AAI.
DR InterPro; IPR001376; Gliadin.
DR InterPro; IPR001954; GIIA_Glutenin.
DR Pfam; PF00234; TRYP_alpha_amy1; 1.
DR PRINTS; PR00208; GLIADGLUTEN.
DR PRINTS; PR00209; GLIADIN.
DR SMART; SMO0499; AAI: 1.
FT CHAIN 1 270 alpha-gliadin.
SQ SEQUENCE 270 AA; 31491 MW; ID84B6528EFA0FF5 CRC64;

Query Match
Best Local Similarity 92.4%; Score 1337.5; DB 2; Length 270;
Matches 251; Conservative 3; Mismatches 12; Indels 1; Gaps 1

QY 1 VRVVPQLOPQNPBQQQPOEQQVPLVQQQQQFFGQQQQQFPPOQYPQOPQPFPSQQPYLQLOP 60
DB 2 VRVSPQLOPQNPBQQQPOEQQVPLVQQQQQFFGQQQQFPFPQYPQLOQPFPSQQPYLQLOP 61
QY 61 FPOEWLPPYQPSQSPFPQYPQOPQYSPQOPISQQQAQQQQQQQQQQQQQ-QCIIQQTL 119
DB 62 FPGQLPFSQQPFRFPQYPQOPQYSPQOPISQQQQQQQQQQQQQQQQQQQQCIIQQTL 121

```







DR EMBL; U51104; AAA96523.1; -.  
DR PIR; S1333; S1333.  
DR PIR; T06282; T06282.  
DR GO; GO:0045755; F:mutant reservoir activity; IEA.  
DR InterPro; IPR003512; AA1.  
DR InterPro; IPR001576; Gliadin.  
DR InterPro; IPR001554; Gliad glutenin.  
DR Pfam; PF00234; TYP\_alpha\_amy; 1.  
DR PRINTS; PR00208; GLIADGUTEN.  
DR PRINTS; PR00209; GLIADIN.  
DR SMART; SM00493; AA1; 1.  
SQ SEQUENCE 268 AA; 3325 MW; D7F6B9133283CA2 CRC64;

Query Match	91.1%;	Score 1309;	DB 2;	Length 288;
Best Local Similarity	92.5%;	Pred. No. 9,8e-73;		
Matches 248;	Conservative 3;	Mismatches 15;	Indels 2;	Gaps 1;

[illegible]

ID	RESULT 15	PRELIMINARY;	PRT;	287 AA.
Q41528				
AC	Q41528			
DT	01-NOV-1996 (TREMBLrel. 01, Created)			
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)			
DE	Alpha-glutinin.			
OS	Triticum aestivum (Wheat).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Magnoliopsida; Liliopsida; Poales; Poaceae			
OC	Triticaceae; Triticum.			
OX	NCBI_Taxid=4565;			
RX	[1]			
SEQUENCE FROM N.A.				
RP	STRAIN=Cheyenne;			
RA	Anderson O.D.;			
RL	Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases			
DR	EMBL; U50964; AAA86276.1;..			
DR	GO; GO:0045735; f:nutrient reservoir activity; IFA.			
DR	InterPro; IPR003612; AAI.			
DR	InterPro; IPR001376; Glia.gluadin.			
DR	InterPro; IPR001954; Glia.gluadin.			
DR	Pfam; PF00234; Tryp.alpha.amyl. 1.			
DR	PRINTS; PR00208; GLIADGLUTEN.			
DR	PRINTS; PR00209; GLIADIN.			
DR	SMART; SM00399; AAI; 1..			
DR	SEQUENCE 287 AA; 33193 MW; 05F82296749C9E87 CRC64;			

Query Match	90.2%	Score 1296.5;	DB 2;	Length 287;
Best Local Similarity	92.9%	Pred. No. 5.7e-72;		
Matches 249; Conservative		2; Mismatches 14;	Indels 3;	Gaps 3;

1 VRIIVPQLQLQNPSQQQPPQEQVPLVQQQQQPPGGQQQQPPPQQQPYPPQPPPPPSQQPPYLQLQP 600

```

Db      21 VRFVVPQQLQONPSCQLPEGEQVPLVQ000QLFGQ0QPFPPQ0QPYQ0PQ- FSQLVYLQ0LP 76
QY      61 FPEPMPLPYQPOSFPFPQ0QPYQPOQ0QYQ0P0PIS- Q0QA0Q00000000000001LQ0L 119
Db      80 FPG0QLPYSPQPFRRQ0QPYQ0PQ0YQ0SPQ0PQISQ00000000000000001LQ0L 139
QY      120 Q00LIFPMQDVLYQ0NHIAHARSQVTLQ0STY0LLQZLCC0HLMQIPESQCCALHNVYHAI 179
Db      140 Q00LIFCMQDVLYQ0NHKAHGRSQVTLQ0STY0LLRRLCC0HLMQIPESQCCALHNVYHAI 199
QY      180 ITHQ00K-0000PSSQVSFQ0PLQ0QYPLQ0G0SFRP0Q0NPOAQ0SVQ0P0QLPFEIIRNL 238
Db      200 ITHQ00K00000PSSQVSFQ0PLQ0QYPLQ0G0SFRP0Q0NPPQ0SVQ0PLPQEEIRNL 259
QY      239 ALQ0TLPMQNVYIAPYCTIAPFGIRGIN 266
Db      260 ALQ0TLPMQNVYIAPYCTIAPFGIRGIN 287

Search completed: December 14, 2004, 17:25:23
Job time : 80 secs

```

Search completed: December 14, 2004, 17:25:23  
Job time : 80 secs

QY 239 ALQTLPMCNVYIAPYCTIAPFGIFGYN 266  
|||:|||||  
Db 260 ALQTLPSMCNVYIPPYCTIAPFGIFGYN 287

Db 200 ILHQQKQKQKQKQSSQVFFQQLQGYPLGGGSRFRPSQGNPQTQGSVQPLPQEEIRNL 259

Db 140 QQLIPCM DVVLQOHNKAHGRS QVLAQSYQLRLRELCCQHLMOIPESQCAIHNVTHAI 199

00 FFQEFQLFISLVYVFNFVUZYFYUYXYYXZ 1  
DB 80 FFFQEFQLFISLVYVFNFVUZYFYUYXYYXZ 1  
OY 120 QQQILPCMDDVLQOHNIAHARSQVTQOSTYOILLQEJLCCGHLMQIPESQCQAHHNVTHAI 179

61 FPGMLPYFGPQSFPPQGPYPQPGQYSGPQGPIS-QQAGQQQQQQQQQQQQQQQILQQL 119

21 VPEEVPOLAPONPSOOLDOEVPPIVQOONFI.GOONPFPEOOPYPOBO-FPSOLPYLOLOP 79

Page 1465-1466



*This Page Blank (uspto)*



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 17:00:05; Search time 76.6667 Seconds  
(without alignments)  
1244.635 Million cell updates/sec

Title: US-10-089-700-3-Y65

Perfect score: 1433  
Sequence: 1 VRVPPQLQPNPSCQQPQE.....CNVTIAPYCTIAPRGIRGTN 266

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: A\_Geneseq23Sep04:\*

- 1: geneseqp1980s:\*
- 2: geneseqp1990s:\*
- 3: geneseqp2000s:\*
- 4: geneseqp2001s:\*
- 5: geneseqp2002s:\*
- 6: geneseqp2003as:\*
- 7: geneseqp2003bs:\*
- 8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1425	99.4	AAU01799	Aau01799 Wheat A-g
2	1425	99.4	ADH14513	Adh14513 A-gliadin
3	1265.5	88.3	AAE38574	Aae38574 Wheat alp
4	1265.5	88.3	ADP19626	Adp19626 Alpha-2-g
5	445	32.4	AAW62647	Aaw62647 Mature-du
6	450.5	31.4	ADOT1669	Adot1669 Amino aci
7	438.5	30.6	ADH89358	Adh89358 T. aestiv
8	437.5	30.5	ADG44134	Adg44134 T. aestiv
9	285.5	19.9	ADJ71661	Adj71661 Amino aci
10	285.5	19.9	ABR63057	Abbr63057 Drosophil
11	266	18.6	AAV96255	Aav96255 Kaposi's
12	266	18.6	AAV85800	Aav85800 HHV8 ORF
13	266	18.6	AAV82331	Aav82331 Amino aci
14	266	18.6	AAV82331	Aav82331 Kaposi's
15	266	18.6	ABR05621	Abbr05621 HHV8 late
16	266	18.6	ADJ50956	Adj50956 Human Pol
17	264.5	18.5	ADOC0905	Adoc0905 Novel hum
18	264.5	18.5	ABO07138	Ab07138 Human nuc
19	264.5	18.5	ADJ37233	Adj37233 Amino aci
20	253.5	17.8	ADOC47673	Adoc47673 H. vulgar
21	253	17.7	ADH89336	Adh89336 H. vulgar
22	253	17.7	ADG44132	Adg44132 S. cerevi
23	253	17.7	ABG93053	Abg93053 S. cerevi
24	253	17.7	ABR53130	Abrr53130 Protein s
25	253	17.7	ADK62564	Adk62564 Disease t

26	247.5	17.3	900	4	ABR62018
27	244.5	17.1	1069	4	ABR61305
28	244	17.0	1013	4	ABR71039
29	239	16.7	358	7	ADH65556
30	236.5	16.5	1142	7	ADOC7968
31	231.5	16.2	149	4	ADH89335
32	231	16.1	160	7	ADH89335
33	231	16.1	160	8	ADG44131
34	231	16.1	160	8	ADH89335
35	230	16.1	2237	5	ABR70004
36	230	16.1	2703	4	ABR60074
37	230	16.1	4365	6	ABU02252
38	228	15.9	158	3	AAV54568
39	228	15.9	2280	4	ABR61650
40	226	15.8	153	3	AAV69495
41	226	15.8	738	5	ABG93140
42	224	15.6	1428	4	ABR70377
43	223.5	15.6	785	8	ADP98983
44	223	15.6	368	4	ABR63167
45	222.5	15.5	467	8	ADJ76333

## ALIGNMENTS

RESULT 1  
AAU01799 standard; protein; 266 AA.  
ID AAU01799

XX AC AAU01799;  
XX DT 07-SEP-2001 (first entry)  
XX DE Wheat A-gliadin.

XX KM Wheat; A-gliadin; epitope; coeliac disease; gluten intolerance;  
XX KW T-cell binding; antagonist; transglutaminase; transgenic plant.

XX OS Triticum aestivum.

XX PN WC000125793-A2.

XX PD 12-APR-2001.

XX PF 02-OCT-2000; 2000WC-GB003760.

XX PR 01-OCT-1999; 99GB-00023306.

XX PA (ISIS-) ISIS INNOVATION LTD.

XX PI Anderson RP, Hill AVS, Jewell DP,

XX WPI; 2001-300179/31.

PT Diagnosing coeliac disease or susceptibility to the disease in an individual, by detecting in vitro or in vivo T cells which bind immunodominant T cell epitope obtained from naturally occurring homolog of gliadin.

PT Claim 1; Page 52; 107p; English.

The sequence represents wheat A-gliadin. A-gliadin derived peptides of the invention are used to test mammalian (preferably human) susceptibility to coeliac disease (gluten intolerance). The peptides are contacted with a blood sample and T cell recognition measured, a positive T-cell recognition indicating a susceptibility to coeliac disease. The peptides are useful for inducing tolerance in an individual and for producing an antibody or preventing coeliac disease in an individual and for producing an antibody specific to the amino acids in length) whose wild-type sequence can be modified by transglutaminase to a sequence that comprises the epitope, but which has been modified in such a way that it does not contain sequence which can



GC be modified by transglutaminase to a sequence that comprise the epitope  
 GC is useful for decreasing the ability of gliadin protein to cause Coeliac  
 GC disease. Nucleic acids encoding proteins antagonistic to the T-cell  
 GC binding of the epitopes are useful for obtaining a transgenic plant cell  
 GC or seed and for the production of a protein. The resultant crop plant is  
 GC useful for obtaining a product of a wheat plant, especially grain, which  
 GC is optionally processed into flour or another grain product. Food  
 GC comprising the antagonistic protein is useful instead of a wild-type  
 CC gliadin

XX Sequence 266 AA;

Query Match 99.4%; Score 1425; DB 4; Length 266;  
 Best Local Similarity 99.6%; Pred. No. 2.1e-115;  
 Matches 265; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VRVVPVQLQPNPSCQCPQEQVPLVQQQPFPGQQQPFPPQPPYPPQPPFSQCPYTLQLP 60  
 DB 1 VRVVPVQLQPNPSCQCPQEQVPLVQQQPFPGQQQPFPPQPPYPPQPPFSQCPYTLQLP 60  
 QY 61 FPGPYLPFPQPSFPQPPQPPYPPQPPYSPQPPISQQAQQQQQQQQQQQQQQQQQQQQ 120  
 DB 61 FPGPYLPFPQPSFPQPPQPPYPPQPPYSPQPPISQQAQQQQQQQQQQQQQQQQQQQQ 120  
 QY 121 QQLIPCMQDVVLQOHNIHARSQVLAQSTYQLLQELCCQHMQLPESQCCAIHNVVAII 180  
 DB 121 QQLIPCMQDVVLQOHNIHARSQVLAQSTYQLLQELCCQHMQLPESQCCAIHNVVAII 180  
 QY 181 LHQQKQKQQQPSQSVSPQPPYPPQPPYSPQPPISQQAQQQQQQQQQQQQQQQQQQQQ 240  
 DB 181 LHQQKQKQQQPSQSVSPQPPYPPQPPYSPQPPISQQAQQQQQQQQQQQQQQQQQQQQ 240  
 QY 241 QTLPMQNVYIAPYCTIAPFGIFGTN 266  
 DB 241 QTLPMQNVYIAPYCTIAPFGIFGTN 266

#### RESULT 2

ADH14513

XX ADH14513;

XX 11-MAR-2004 (first entry)

XX A-gliadin protein sequence SEQ ID NO:3.

XX coeliac disease; gliadin; gliadin T cell epitope; gastrointestinal;

XX vaccine.

XX Synthetic.

XX MO2003104273-A2.

XX 18-DEC-2003.

XX 05-JUN-2003; 2003MO-GB002450.

XX 05-JUN-2002; 2002GB-00012885.

XX (ISIS-) ISIS INNOVATION LTD.

XX Anderson RP, Hill AVS, Jewell DP;

XX WPT; 2004-043640/04.

XX Preventing or treating coeliac disease comprises administering agent

XX which are wheat gliadin T cell epitope capable of being recognized by T

XX cell receptor.

XX Example 1; SEQ ID NO 3; 177pp; English.

XX The present invention describes a method (M1) for preventing or treating

CC coeliac disease. M1 comprises administering an agent (A) comprising a  
 CC gliadin T cell epitope, which is capable of being recognised by a T cell  
 CC receptor, to an individual. Gliadin is a component of gluten. (A) has  
 CC gastrointestinal activity, and can be used in vaccines. The agent (A) can  
 CC be used in the preparation of a medicament for treating or preventing  
 CC coeliac disease. (A) can also be used in the preparation of a diagnostic  
 CC means for use in diagnosing coeliac disease, or susceptibility to coeliac  
 CC disease, in an individual, which involves determining whether T cells of  
 CC the individual recognise the agent, recognition by the T cells indicating  
 CC that the individual has, or is susceptible to, coeliac disease. The  
 CC present sequence represents a protein which is used in the  
 CC exemplification of the present invention.

XX Sequence 266 AA;

Query Match 99.4%; Score 1425; DB 8; Length 266;  
 Best Local Similarity 99.6%; Pred. No. 2.1e-115;  
 Matches 265; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VRVVPVQLQPNPSCQCPQEQVPLVQQQPFPGQQQPFPPQPPYPPQPPFSQCPYTLQLP 60  
 DB 1 VRVVPVQLQPNPSCQCPQEQVPLVQQQPFPGQQQPFPPQPPYPPQPPFSQCPYTLQLP 60  
 QY 61 FPGPYLPFPQPSFPQPPQPPYPPQPPYSPQPPISQQAQQQQQQQQQQQQQQQQQQQQ 120  
 DB 61 FPGPYLPFPQPSFPQPPQPPYPPQPPYSPQPPISQQAQQQQQQQQQQQQQQQQQQQQ 120  
 QY 121 QQLIPCMQDVVLQOHNIHARSQVLAQSTYQLLQELCCQHMQLPESQCCAIHNVVAII 180  
 DB 121 QQLIPCMQDVVLQOHNIHARSQVLAQSTYQLLQELCCQHMQLPESQCCAIHNVVAII 180  
 QY 181 LHQQKQKQQQPSQSVSPQPPYPPQPPYSPQPPISQQAQQQQQQQQQQQQQQQQQQQQ 240  
 DB 181 LHQQKQKQQQPSQSVSPQPPYPPQPPYSPQPPISQQAQQQQQQQQQQQQQQQQQQQQ 240  
 QY 241 QTLPMQNVYIAPYCTIAPFGIFGTN 266  
 DB 241 QTLPMQNVYIAPYCTIAPFGIFGTN 266

#### RESULT 3

AAE38574

XX AAE38574;

XX 04-DEC-2003 (first entry)

XX Wheat alpha-2 gliadin protein.

XX Wheat; therapy; celiac sprue; dermatitis herpetiformis; gluten toxicity;

XX glutenase; foodstuff; antiinflammatory; dermatological; alpha-2 gliadin.

XX Triticum aestivum.

XX WO2003068170-A2.

XX 21-AUG-2003.

XX 14-FEB-2003; 2003WO-US004743.

XX 14-FEB-2002; 2002US-0357238P.

XX 14-MAY-2002; 2002US-0380761P.

XX 28-JUN-2002; 2002US-0392782P.

XX 31-OCT-2002; 2002US-0432933P.

XX 20-NOV-2002; 2002US-0458033P.

XX 20-DEC-2002; 2002US-045881P.

XX (STRD ) UNIV LEIAND STANFORD JUNIOR.

XX Hausch F, Gray G, Shan L, Khosla C;

XX WPT; 2003-697466/66.







XX The present sequence represents the mature glutenin protein. The DNA  
 CC sequence encoding this protein is isolated from the genomic DNA of  
 CC Triticum durum L. The gene codes for a low-molecular-weight glutenin  
 CC protein and can be used to produce transgenic durum wheat plants with  
 CC "better quality characteristics" (no details given). (Updated on 25-MAR-  
 CC 2003 to correct PI field.) (Updated on 17-OCT-2003 to standardise OS  
 CC field)  
 XX  
 XX Sequence 369 AA;  
 SO  
 Query Match 32.4%; Score 465; DB 2; Length 369;  
 Best Local Similarity 40.2%; Pred. No. 5,6e-32;  
 Matches 134; Conservative 32; Mismatches 87; Indels 80; Gaps 14;  
 6 PQLPQNPSSQPP-----QEQVPLVQGGQPPGQGGQPP-----PQPPVQPPPP----- 50  
 38 PQGQPCGQGGQPPPLSQGGQPPPSQGGQPPPSQGGQPPVQLQGSFSGQGLPPPSQGGQPP 97  
 51 ---SQQPYL-----QLQPPQPYLPYPPQPSPPPP-----PYPPQPYGQ 83  
 98 FSGQGGQPPVLPQPPPSQGGQPPPSQGLPPPSQGLPPPSQGGQPPPSQGLPPPSQ 157  
 90 PQGQPCGQGGQPPPLSQGGQPPPSQGGQPPVQLQGGQPPVQLQGGQPPVQLQGGQPP 132  
 158 QGQVPLVQGGQPPPSQGGQPPPSQGGQPPVQLQGGQPPVQLQGGQPPVQLQGGQPP 216  
 133 QH-----NIAHARSQVLAQSTYQLQGLCCQHLMOIPEDSGCCQALHNVAIILHQOK 186  
 217 QGQSPWAMPQSLASSQMLQSSCHVMQGGQGLPQIPQSSRYEALRAIVYSIIL--QEQ 274  
 187 QGQSPSSQVSPQGLQGYPRGQSGFRPSQGNPQAGS-----VQPPQLPPFEE 234  
 275 QGQVGSQGGQGGQGGQGGQGGQGGQGGQGGQGGQGGQGGQGGQGGQGGQGGQGGQGG 311  
 235 IRLNALQTLPAMCNVYIAPY--CTIAPFGIFGT 265  
 332 MTSIALRTPLTMCNWNVIVRTTRVDFGV-GT 363  
 DB  
 QY 235 IRLNALQTLPAMCNVYIAPY--CTIAPFGIFGT 265  
 DB 332 MTSIALRTPLTMCNWNVIVRTTRVDFGV-GT 363  
 RESULT 6  
 ADO71669  
 ID ADO71669 standard; protein; 297 AA.  
 XX  
 AG ADO71669;  
 XX  
 DE 12-AUG-2004 (first entry)  
 XX  
 DE Amino acid sequence of a modified glutenin LMW subunit.  
 KW low molecular weight subunit; LMW subunit; glutenin;  
 KW wheat cultivar Cheyenne; gliadin; flour; tablet; coeliac disease;  
 KW gluten intolerance.  
 XX  
 OS Triticum sp.  
 OS Synthetic.  
 XX  
 EN BP1424342-A1.  
 XX  
 DE 02-JUN-2004.  
 XX  
 DE 27-NOV-2002; 2002EP-00026461.  
 XX  
 DE 27-NOV-2002; 2002EP-00026461.  
 XX  
 DE (BAKE-) BAKEMARK DEUT GMBH.  
 DE (MONS-) MONSANTO AGRAR DEUT GMBH.  
 DE (UNIF-) UNIFERN GMBH & CO KG.  
 DE (PURA-) PURATOS NV.  
 XX  
 DE Hinzmann E, Wieser H, Stahl U;  
 DE MPI; 2004-402870/38.

DR N-PSDB; ADO71668.  
 XX  
 XX Novel nucleic acid comprising sequence encoding modified glutenin  
 PT polypeptide, useful for preparing modified glutenin polypeptide as  
 PT gliadin substitute in foodstuffs such as dough, pastries and wafers.  
 XX  
 XX Claim 16; Fig 11; 43pp; English.  
 XX  
 CC The present sequence represents a modified low molecular weight (LMW)  
 CC subunit of glutenin. The wild type subunit is designated clone LMW6, and  
 CC is isolated from wheat cultivar Cheyenne. The LMW6 polypeptide does not  
 CC contain the allergenic epitope GQGP, and shows some minor differences to  
 CC published sequences. It therefore represents a new allele for LMW subunit  
 CC genes. The LMW6 polypeptide was modified to produce modified glutenin  
 CC polypeptides of the invention. In these modified polypeptides one or more  
 CC cysteine residues responsible for intermolecular cross linking through  
 CC disulfide bridges are deleted or substituted. The modified glutenin  
 CC polypeptide is useful as a gliadin substitute. It is also useful in the  
 CC preparation of foodstuffs, such as flour or for the preparation of  
 CC pharmaceutical products, such as tablets, where the foodstuffs contain a  
 CC considerably reduced amount of gliadin proteins or no gliadin proteins.  
 CC Pharmaceutical compositions comprising the modified polypeptide of the  
 CC invention are useful for treating patients suffering from coeliac disease  
 CC or persons who are intolerant to gluten.  
 XX  
 XX Sequence 297 AA;  
 SO  
 Query Match 31.4%; Score 450.5; DB 8; Length 297;  
 Best Local Similarity 41.2%; Pred. No. 7.9e-31;  
 Matches 120; Conservative 38; Mismatches 82; Indels 51; Gaps 13;  
 5 VPQLPQNPSSQGGQPPGQGGQPPGQGGQPPGQGGQPPGQGGQPPGQGGQPPGQGGQPP 63  
 18 IQMETSIPGSEFWQGGQPPGQGGQPPGQGGQPPGQGGQPPGQGGQPPGQGGQPPGQGGQPP 71  
 64 PVLPPQPSPP 116  
 72 -----GLFSGQKQPPVLPQGPAPPSQGGQGLPQGPAPPSQGGQGLPQGPAPPSQGGQ 125  
 117 QI-----LQGGQPLPMDVVLQGNINIAHARSQVLAQSTYQLQGLCCQHLMOIPEDSGCC 170  
 126 QLNPKVFLQGGQ-----GSPVAMPQH--LASSQMWQSSCNVMQGGQGGQGLPRLPESQRYE 179  
 171 AIHNVAIILHQOK-----QGGQPSQV--SFQGPVQ--QYPLGQGSF-----RP 213  
 180 AIPRAIFSIILQEGQGGQGGQGGQGGQGGQGGQGGQGGQGGQGGQGGQGGQGGQGGQGGQ 239  
 214 SQGNPQAGSVQPPQGLPQGGQGGQGGQGGQGGQGGQGGQGGQGGQGGQGGQGGQGGQGG 262  
 240 QGQVQKGTFLQPPHIALRLEVMTSIALRTPLTMCNWNVIVYSITSPAGV 290  
 DB  
 QY 214 SQGNPQAGSVQPPQGLPQGGQGGQGGQGGQGGQGGQGGQGGQGGQGGQGGQGGQGGQGG 262  
 DB 240 QGQVQKGTFLQPPHIALRLEVMTSIALRTPLTMCNWNVIVYSITSPAGV 290  
 RESULT 7  
 ADH89338  
 ID ADH89338 standard; protein; 307 AA.  
 XX  
 AC ADH89338;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DT T. aestivum LMW glutenin-1D1 protein.  
 XX  
 DE double stranded RNA; storage protein; 2S-albumen; 7S-globulin;  
 KW 11S/12S-globulin; zein-prolamine; homogenistate metabolic pathway;  
 KW pharmaceutical; plant; abiotic stress; fatty acid composition;  
 KW lipid composition; oil composition; carbohydrate composition; colour;  
 KW pigmentation; pathogen resistance; fruit ripening delay; aging;  
 KW male sterility; lignin; fibre; cotton; Vitamin E synthesis; nicotine;  
 KW caffeine; theophylline; threonine biosynthesis; glutenin.  
 XX  
 OS Triticum aestivum.  
 OS  
 PN WO2003078629-A1.







Query	Subject	Score	Expect	Ident	Length	Positives	Negatives	Gaps	Mismatches	Indels
1	1	30.5%	437.5	DB 8	298					
2	2	30.5%	437.5	DB 8	298					
3	3	30.5%	437.5	DB 8	298					
4	4	30.5%	437.5	DB 8	298					
5	5	30.5%	437.5	DB 8	298					
6	6	30.5%	437.5	DB 8	298					
7	7	30.5%	437.5	DB 8	298					
8	8	30.5%	437.5	DB 8	298					
9	9	30.5%	437.5	DB 8	298					
10	10	30.5%	437.5	DB 8	298					
11	11	30.5%	437.5	DB 8	298					
12	12	30.5%	437.5	DB 8	298					
13	13	30.5%	437.5	DB 8	298					
14	14	30.5%	437.5	DB 8	298					
15	15	30.5%	437.5	DB 8	298					
16	16	30.5%	437.5	DB 8	298					
17	17	30.5%	437.5	DB 8	298					
18	18	30.5%	437.5	DB 8	298					
19	19	30.5%	437.5	DB 8	298					
20	20	30.5%	437.5	DB 8	298					
21	21	30.5%	437.5	DB 8	298					
22	22	30.5%	437.5	DB 8	298					
23	23	30.5%	437.5	DB 8	298					
24	24	30.5%	437.5	DB 8	298					
25	25	30.5%	437.5	DB 8	298					
26	26	30.5%	437.5	DB 8	298					
27	27	30.5%	437.5	DB 8	298					
28	28	30.5%	437.5	DB 8	298					
29	29	30.5%	437.5	DB 8	298					
30	30	30.5%	437.5	DB 8	298					
31	31	30.5%	437.5	DB 8	298					
32	32	30.5%	437.5	DB 8	298					
33	33	30.5%	437.5	DB 8	298					
34	34	30.5%	437.5	DB 8	298					
35	35	30.5%	437.5	DB 8	298					
36	36	30.5%	437.5	DB 8	298					
37	37	30.5%	437.5	DB 8	298					
38	38	30.5%	437.5	DB 8	298					
39	39	30.5%	437.5	DB 8	298					
40	40	30.5%	437.5	DB 8	298					
41	41	30.5%	437.5	DB 8	298					
42	42	30.5%	437.5	DB 8	298					
43	43	30.5%	437.5	DB 8	298					
44	44	30.5%	437.5	DB 8	298					
45	45	30.5%	437.5	DB 8	298					
46	46	30.5%	437.5	DB 8	298					
47	47	30.5%	437.5	DB 8	298					
48	48	30.5%	437.5	DB 8	298					
49	49	30.5%	437.5	DB 8	298					
50	50	30.5%	437.5	DB 8	298					
51	51	30.5%	437.5	DB 8	298					
52	52	30.5%	437.5	DB 8	298					
53	53	30.5%	437.5	DB 8	298					
54	54	30.5%	437.5	DB 8	2					



DR N-PSDB; ABL15798.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signaling and cell-cell  
PT interactions.  
XX  
PS Disclosure; SEQ ID NO 41877; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signaling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (AB57737-  
CC AB872072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 1798 AA;  
XX  
Query Match 19.9%; Score 285.5; DB 4; Length 1798;  
Best Local Similarity 37.3%; Pred. No. 1.3e-15;  
Matches 100; Conservative 18; Mismatches 93; Indels 57; Gaps 12;  
XX  
QY 3 VPVPQLQPNP-SQQQPEQVPLVQ-----QQQPPQQQQPPPPQ 41  
DB 264 VEGQATQPFQSFQCKFIDPTDPCVAVLSALSNDLIMRQQLKQQQDMQQQQQ 323  
QY 42 PYPQPPFP-SQQPVLQCPFPQFVLPYRQPSFPQQ-----PYPQPPQVQSPQPPISQ 96  
DB 324 WAPQPPQQAQPPQQQQQQQQPPQQ-QHTSPQSPFQQQPTPTPTLQQQPNQQAQ-1QQ 381  
QY 97 QQAQQQQQQQQQQQQQQQQQLLQQLLQQLLP-CMDVVLQGH-NIAHARSQVLQGSTYQL 152  
DB 382 QQQQQQQQQQQQQQQQQQQQVLTQQQPPQQQQQVLTGRHVINTSTAGCQQLIQSHMSL- 440  
QY 153 QSLCCGHLWQIPBQSCCAIHNVVAILTHQQKQKQSSQVSPQPLQVPLQGSFR 212  
DB 441 -----ALQVQQQ-----LTHVQQQAQQAQPPQQQQITTVQQLPFAQQQQQL 480  
QY 213 PQQQNPQAQSVQPPQLPQFEETRNAL 240  
DB 481 PQGHVQDQ---QPPQV-QFTQQQQIAL 504  
XX  
RESULT 11  
ABB63057  
ID ABB63057 standard; protein; 2285 AA.  
XX  
AC ABB63057;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polyprotein S80 ID NO 15963.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO2001.71042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PE 23-MAR-2001; 2001WO-US009231.  
XX  
PR 23-MAR-2000; 2000US-0191637P.  
XX  
PR 11-JUL-2000; 2000US-00614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
FI Venter JC, Adams N, Li FMD, Myers EM;

XX  
DR WPI; 2001-656860/75.  
DR N-PSDB; ABL07160.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signaling and cell-cell  
PT interactions.  
XX  
PS Disclosure; SEQ ID NO 15963; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signaling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins (AB57737-  
CC AB872072). The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 2285 AA;  
XX  
Query Match 18.8%; Score 269; DB 4; Length 2285;  
Best Local Similarity 35.7%; Pred. No. 4.5e-14;  
Matches 101; Conservative 14; Mismatches 104; Indels 64; Gaps 13;  
XX  
QY 7 QLPQNP-----SQQPEQVPLVQQQFPQQQQFPQPPQPPQ-----PFP-SQQP 54  
DB 909 QMQQQQVPAVPAVTHVQVMPQQQVN-QQQQQQPMQQLPQQVQVQPPVLPFPQHQ 966  
QY 55 YLQHPFPQ-----PYLPYRQPSFPQQFVYPQPPQVQSPQPPISQQA--- 99  
DB 967 QQQQQPQLQLQLMHTNVQAPL--TQQQVMAQQQAQVYFQQQQQQQPPQAVNMQQAAYAM 1024  
QY 100 QQQQQQQQQQQQQQQQLLQQLLQQLI-----PCMDVVLQGHNTAHARSQVLQGSTYQL 153  
DB 1025 QQAQQQQQLSLQPLDQQLLQQQQVAVSHQQQIMQQQLAQHQLQQLQQQQLQQQQ 1084  
QY 154 ELCCGHLWQIPBQSCCAIHNVVAILTHQQKQKQSSQV-----SFQQPLQ-----QY 204  
DB 1085 QIQQQQLQQQQLQQQ-QFVQVYACA--MPQQHQDLVTSQVWAFHQHQPPIDQPVQM 1141  
QY 205 PL-----GQGSFPPSQNPQAQGSVPQQLPQF 232  
DB 1142 PPTSVAPPIDHTYVQCGGVTLSDAQQQQHPPGFAVQQAAPF 1184  
XX  
RESULT 12  
AAV96255  
ID AAV96255 standard; protein; 1162 AA.  
XX  
AC AAV96255;  
XX  
DT 12-SEP-2003 (revised)  
DT 11-SEP-2000 (first entry)  
XX  
DE Kaposi's sarcoma-associated herpesvirus LANA.  
XX  
KW Kaposi's sarcoma-associated herpesvirus; KSHV; rhadino virus;  
KW latency-associated nuclear antigen; LANA; gamma-2 herpes virus;  
KW Human herpes virus 8; HHV8; rhadino virus cis-acting element; RVCAE;  
KW Kaposi's sarcoma; primary effusion lymphoma; PEL;  
KW human immunodeficiency virus; HIV; multicentric Castleman's disease.  
XX  
OS Human herpesvirus 8.  
XX  
FH Key Location/Qualifiers  
FT Domain 14..17  
FT /note= "nuclear localisation signal, NLS"  
FT Domain 64..70  
FT /note= "nuclear localisation signal, NLS"  
FT Region 320..429







```

QY      4 PVPQLQPNPSQQQPOEQLVTVQ--QQQFPGQQ--QQFPPQQPYPPQPPSPQQPYLQLO 59
DB      495 PLQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQ 553
QY      60 P---PQQ---PYLPYQQ---PQSFPQQPYPPQ--QPYSSQPQQPISQQQAQQQQQQQQQQ 109
DB      554 PQQREPPQQREPPQQREPPQQREPPQQREPPQQREPPQQREPPQQREPPQQREPPQQRE 613
QY      110 QQQQQLLQQLLQQLLPCMDVVLQGHNTAHARSQVLQGSTYQLLQELCCGHLMQIPRQSC 169
DB      614 EQQQDEQQQDEQQQ---QDEQQQDEQQQDEQQQDEQQQDEQQQDE---QQQDEQQQD 663
QY      170 QALHNVVHAILHQQKQQQPSQVSPFQQPLQYPLGGSPFRPSQGNPQAQGSVPQQL 229
DB      664 EQQQD-----EQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQD 716
QY      230 PQFEE 234
DB      717 EQQQD 721

```

## RESULT 14

AAB62331  
ID AAB62331 standard; protein; 1162 AA.

AC AAB62331;

DT 06-AUG-2003 (revised)  
DT 29-JUN-2001 (first entry)

DE Amino acid sequence of KSHV tethering protein LANA.

KW Histone H1; tethering protein; LANA; gene therapy; multiple sclerosis;  
KW Parkinson's disease; Huntington disease; diabetes; human herpesvirus 8;  
KW KSHV; latency-associated nuclear antigen; LANA.

OS Human herpesvirus 8.

PN WO200125484-A2.

PD 12-APR-2001.

PF 29-SEP-2000; 2000WO-US026908.

PR 01-OCT-1999; 99US-00410399.

PA (UNMI ) UNIV MICHIGAN.

PI Robertson BS, Cotter MA;

DR WPI; 2001-281736/29.

XX N-PDB; AAF82901.

PT A composition for use in gene therapy comprises an expression vector that  
PT includes a nucleic acid sequence encoding a nucleic acid binding protein.

PS Disclosure; Fig 9B; 60pp; English.

XX The invention provides a composition comprising nucleic acid, histone H1  
CC protein and expression vector operationally encoding a protein suitable  
CC for tethering the nucleic acid to the histone H1 protein, where the  
CC tethering protein is LANA. The composition is useful in adding the  
CC retention of the viral DNA in the host cell. The viral vector encodes a  
CC protein suitable for tethering DNA to Histone H1. Methods for screening of  
CC for compounds which are agonistic or antagonistic for the tethering of  
CC viral proteins to histone H1 and DNA binding sites are useful for  
CC developing the method of viral transfer. The composition has applications  
CC to gene therapy, including the treatment of multiple sclerosis,  
CC Parkinson's disease, Huntington disease and diabetes. The present  
CC sequence represents the amino acid sequence of the Kaposi's sarcoma  
CC associated herpesvirus (human herpesvirus 8) latency-associated nuclear  
CC antigen (LANA), which acts as a tethering protein. (Updated on 06-AUG-

CC 2003 to correct OS field.)

XX SQ Sequence 1162 AA;

Query Match 18.6%; Score 266; DB 4; Length 1162;  
Best Local Similarity 38.0%; Fred. No. 3, 8e-14;  
Matches 93; Conservative 19; Mismatches 101; Indels 32; Gaps 10;

```

QY      4 PVPQLQPNPSQQQPOEQLVTVQ--QQQFPGQQ--QQFPPQQPYPPQPPSPQQPYLQLO 59
DB      495 PLQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQEPQQQ 553
QY      60 P---PQQ---PYLPYQQ---PQSFPQQPYPPQ--QPYSSQPQQPISQQQAQQQQQQQQQ 109
DB      554 PQQREPPQQREPPQQREPPQQREPPQQREPPQQREPPQQREPPQQREPPQQREPPQQRE 613
QY      110 QQQQQLLQQLLQQLLPCMDVVLQGHNTAHARSQVLQGSTYQLLQELCCGHLMQIPRQSC 169
DB      614 EQQQDEQQQDEQQQ---QDEQQQDEQQQDEQQQDEQQQDEQQQDE---QQQDEQQQD 663
QY      170 QALHNVVHAILHQQKQQQPSQVSPFQQPLQYPLGGSPFRPSQGNPQAQGSVPQQL 229
DB      664 EQQQD-----EQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQD 716
QY      230 PQFEE 234
DB      717 EQQQD 721

```

## RESULT 15

ABB05621  
ID ABB05621 standard; protein; 1162 AA.

AC ABB05621;

DT 25-APR-2002 (first entry)

DE Kaposi's sarcoma-associated herpesvirus LANA protein.

KW Kaposi's sarcoma-associated herpesvirus; KSHV; LANA; RVCAE; PEL;  
KW KSHV terminal repeat; rhadino virus cis acting element; episome;  
KW primary effusion lymphoma; latency-associated nuclear antigen;  
KW gene therapy; gene transfer.

OS Human herpesvirus 8.

PN US6322792-B1.

PD 27-NOV-2001.

PF 21-APR-1999; 99US-00298568.

PR 19-NOV-1998; 98US-0109422P.

PA (KIEF/) KIEFF E D.

PI Kieff ED, Ballestas ME, Kaye KW;

DR WPI; 2002-153769/20.

XX N-PDB; ABA93487.

PT System for episomal retention of plasmids in mammalian cells, useful in  
PT gene therapy, comprises rhadinoviral LANA and RVCAE sequences.

PS Disclosure; Fig 7; 27pp; English.

XX The present invention describes a system (A) for maintaining a plasmid as  
CC an episome in mammalian cells, comprising the rhadinoviral sequence LANA  
CC (latency-associated nuclear antigen) of 3489 base pairs (see ABA93487,  
CC 51) expressed in the cell, and the rhadinoviral sequence RVCAE  
CC (rhadinoviral cis-acting element) of 801 base pairs (see ABA93488, 52)  
CC present in the plasmid. Also describes is a method for maintaining a  
CC closed circular DNA in a cell by expressing (51) in the cells and having







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 14, 2004, 17:06:20 ; Search time 19 Seconds

(without alignments)  
928.452 Million cell updates/sec

Title: US-10-089-700-3-Y65

Perfect score: 1433

Sequence: 1 VRVVPQLQPNPQSQDQFQ.....CNVYIAPYCTIAPGIRGCTN 266

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*

2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*

3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*

4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*

5: /cgn2\_6/ptodata/1/1aa/PCTUS.COMB.pep:\*

6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	465	32.4	369	2	US-08-991-300-2
2	266	18.6	1162	2	US-08-728-323A-2
3	266	18.6	1162	3	US-09-298-568-2
4	266	18.6	1162	4	US-09-410-399-2
5	266	18.6	1162	4	US-09-894-273-2
6	267	17.2	788	4	US-08-918-914-4
7	238	16.6	498	4	US-09-270-767-45042
8	234	16.3	256	4	US-09-248-796A-21251
9	219	15.3	2074	4	US-09-491-356C-9
10	218	15.2	579	4	US-09-668-119-3
11	201	14.0	2023	4	US-09-491-356C-8
12	201	14.0	2124	4	US-09-538-092-1377
13	199	13.9	663	4	US-09-270-767-61220
14	199	13.9	1591	4	US-09-270-767-45698
15	198.5	13.9	505	4	US-09-248-796A-19253
16	197	13.7	379	4	US-09-248-796A-21759
17	196.5	13.7	2441	1	US-08-194-468-2
18	196.5	13.7	2441	1	US-08-961-739-2
19	196.5	13.7	2441	3	US-09-514-247A-8
20	196.5	13.7	2441	4	US-09-686-316-2
21	196	13.7	216	4	US-09-248-796A-21017
22	195.5	13.6	2442	3	US-09-514-247A-10
23	195.5	13.6	2442	4	US-09-538-092-1370
24	194.5	13.6	729	4	US-09-625-188-20
25	186	13.0	295	4	US-09-248-796A-20004
26	185.5	12.9	320	4	US-09-248-796A-24758
27	184.5	12.9	332	4	US-09-248-796A-21649

## ALIGNMENTS

28	183.5	12.8	316	4	US-09-270-767-42663	Sequence 42663, A
29	183	12.8	169	4	US-09-248-796A-28087	Sequence 28087, A
30	183	12.8	519	4	US-09-248-796A-19263	Sequence 19263, A
31	178	12.4	383	4	US-09-248-796A-23236	Sequence 23236, A
32	178	12.4	408	4	US-09-248-796A-14480	Sequence 14480, A
33	178	12.4	848	4	US-09-538-092-33	Sequence 33, Appl
34	177	12.4	261	4	US-09-602-565-34	Sequence 34, Appl
35	177	12.4	1319	4	US-09-538-092-1281	Sequence 1291, Ap
36	176	12.3	618	4	US-09-248-796A-15319	Sequence 15319, A
37	175.5	12.2	684	4	US-09-823-240A-9	Sequence 9, Appl
38	175	12.2	657	4	US-09-248-796A-19232	Sequence 19232, A
39	174.5	12.2	382	4	US-09-248-796A-18720	Sequence 18720, A
40	174	12.1	1507	4	US-09-914-259-37	Sequence 37, Appl
41	173.5	12.1	675	4	US-09-248-796A-20699	Sequence 20699, A
42	172	12.0	542	1	US-07-814-964-13	Sequence 13, Appl
43	172	12.0	542	1	US-08-258-442-13	Sequence 13, Appl
44	172	12.0	542	1	US-08-328-809-8	Sequence 8, Appl
45	172	12.0	542	4	US-08-866-840-8	Sequence 8, Appl

RESULT 1  
US-08-991-300-2  
Sequence 2, Application US/08991300  
Patent No. 5973225  
GENERAL INFORMATION:  
APPLICANT: D'OVIDIO, RENATO  
APPLICANT: PORCEDDU, ENRICO  
APPLICANT: MERCHETTI, CINZIA  
APPLICANT: CARRELLI, LUISA ERCOLI  
TITLE OF INVENTION: ISOLATION AND CHARACTERIZATION OF A GENE  
ENCODING A LOW MOLECULAR WEIGHT GLUTENIN  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY  
CITY: ARLINGTON  
STATE: VA  
COUNTRY: USA  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/991,300  
FILING DATE: 16-DEC-1997  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: IT MI 96/A 002663  
FILING DATE: 19-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NORMAN F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 2264-0201-0X  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-413-3000  
TELEFAX: 703-413-2220  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 369 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-991-300-2  
Query Match 32.4%; Score 465; DB 2; Length 369;  
Best Local Similarity 40.2%; Pred. No. 1.6e-35;



Matches 134; Conservative 32; Mismatches 87; Indels 80; Gaps 14;

```

QY 6 PQLQPNPSPQPP-----QEVPLVQGGQPGGQQQPP--PQGPYPQPPPP----- 50
DB 38 PQQQPCGQQQQPPLSQQQQPPPSQQQQPPPSQQQQPVLPPQPSFQQQLPPPSQQQQPP 97
QY 51 ---SQEYL-----QLPFPQPLPYPPQSPFPQQ-----PYQPQPPQYSQ 89
DB 98 FQQQQQVPLPQQPFSQQQLPPFSQQQLPPFSQQQLPPFSQQQLPPFSQQQLPPFSQ 157
QY 90 PQQPISQQQAQQQQQQQ-----QQQQQQQLIQQ-----LIQQQLPQCMVVLQ 132
DB 158 QQQQVPLPQQPFSQQQQQPPPPQQPPFSQQQQPVLQQQLPVPHPSLQQQLNC-KVFLQ 216
QY 133 QH-----NTAHASQVLTQSTYQLQLQELCCQHLWQIPESQCCQAHNVVHAITLHQK 186
DB 217 QQQSPWAMPQSIARSGQLQQSSGCVVQQCCQQLPQIFQSSRYEALRAIVYSILL--QEQ 274
QY 187 QQQQSSQVSPQQLQGYPLGQSSFRSQQQPQAQGS-----VQPPQLPQFEE 234
DB 275 QQQVGSITQTOQQQPPQ---LQQCVSPQQQSQQQLGQQPQQQLAHGTLQPHQIAQLEV 331
QY 235 IRLALQTLFAMCNVYIAPY--CTIAPFGIFGT 265
DB 332 MTSIALKTLPTMCNWNVPLVTRTVPEFV-GT 363

```

## RESULT 2

US-08-728-323A-2  
Sequence 2, Application US/08728323A  
Patent No. 5948676

## GENERAL INFORMATION:

APPLICANT: Chang, Yuan  
APPLICANT: Bohenzky, Roy A.  
APPLICANT: Russo, James J.  
APPLICANT: Edelman, Isidore S.  
APPLICANT: Moore, Patrick S.  
TITLE OF INVENTION: Immediate Early Protein From Kaposi's  
TITLE OF INVENTION: Sarcoma-Associated Herpesvirus' DNA  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/728,323A  
FILING DATE:  
CLASSIFICATION: 435

## ATTORNEY/AGENT INFORMATION:

NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 0575/52268/JPM/MSC/SKS  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-278-0400  
TELEFAX: 212-391-0525  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1162 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-728-323A-2

## Query Match

18.6%; Score 266; DB 2; Length 1162;

Best Local Similarity 39.0%; Pred. No. 2,3e-16;  
Matches 93; Conservative 19; Mismatches 101; Indels 32; Gaps 10;

```

QY 4 PVPQLQPNPSPQPPQEVPLVQ---QQQPPGQ--QQFPQQPYPPQPPPSQQPYLQLQ 59
DB 495 PLEPQQQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 553
QY 60 P---FPQ---PYLPYPC---PQSPFPQPPYPP--QPPYSQPPQPSQQQAQQQQQQQQ 109
DB 554 PQQREPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 613
QY 110 QQQQILQQQLIFPCMDVVLQQHNTAHASQVLTQSTYQLQLQELCCQHLWQIPESQSC 169
DB 614 EQQDEQQDEQQ---QDEQQQDEQQQDEQQQDEQQQDEQQQDE---QQDEQQQD 663
QY 170 QAHNVVHAITLHQQKQKQQQSSQVSPQQLQGYPLGQSSFRSQQQPQAQGSVPQQL 229
DB 664 EQQDQ-----EQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQD 716
QY 230 PQFEE 234
DB 717 EQQDQ 721

```

## RESULT 3

US-09-298-568-2  
Sequence 2, Application US/09298568  
Patent No. 6322792

## GENERAL INFORMATION:

APPLICANT: Kleff, Elliott D.  
APPLICANT: Ballestas, Mary E.  
APPLICANT: Kaye, Kenneth M.  
TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO  
TITLE OF INVENTION: VIRUS DNA TO MEDIATE EFFICIENT EPISOME PERSISTENCE  
FILE REFERENCE: 16412-10001R  
CURRENT APPLICATION NUMBER: US/09/298,568  
CURRENT FILING DATE: 1999-04-21  
EARLIER APPLICATION NUMBER: US 60/109,422  
EARLIER FILING DATE: 1998-11-19  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 2  
TYPE: PRT  
ORGANISM: Kaposi's sarcoma-associated herpesvirus  
US-09-298-568-2

Query Match 18.6%; Score 266; DB 3; Length 1162;  
Best Local Similarity 38.0%; Pred. No. 2,3e-16;  
Matches 93; Conservative 19; Mismatches 101; Indels 32; Gaps 10;

```

QY 4 PVPQLQPNPSPQPPQEVPLVQ---QQQPPGQ--QQFPQQPYPPQPPPSQQPYLQLQ 59
DB 495 PLEPQQQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 553
QY 60 P---FPQ---PYLPYPC---PQSPFPQPPYPP--QPPYSQPPQPSQQQAQQQQQQQQ 109
DB 554 PQQREPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 613
QY 110 QQQQILQQQLIFPCMDVVLQQHNTAHASQVLTQSTYQLQLQELCCQHLWQIPESQSC 169
DB 614 EQQDEQQDEQQ---QDEQQQDEQQQDEQQQDEQQQDEQQQDE---QQDEQQQD 663
QY 170 QAHNVVHAITLHQQKQKQQQSSQVSPQQLQGYPLGQSSFRSQQQPQAQGSVPQQL 229
DB 664 EQQDQ-----EQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQQDEQQD 716
QY 230 PQFEE 234
DB 717 EQQDQ 721

```

## RESULT 4



US-09-410-399-2  
; Sequence 2, Application US/09410399  
; Patent No. 6482587  
; GENERAL INFORMATION:  
; APPLICANT: Robertson, Etle S.  
; APPLICANT: Colter, Murray A.  
; TITLE OF INVENTION: Methods to Inhibit or Enhance the Binding of Viral DNA  
; FILE REFERENCE: UM-03778  
; CURRENT APPLICATION NUMBER: US/09/410,399  
; CURRENT FILING DATE: 1999-10-01  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1162  
; TYPE: PRT  
; ORGANISM: Kaposi's sarcoma-associated herpesvirus  
US-09-410-399-2

Query Match 18.6%; Score 266; DB 4; Length 1162;  
Best Local Similarity 38.0%; Pred. No. 2.3e-16;  
Matches 93; Conservative 19; Mismatches 101; Indels 32; Gaps 10;

QY 4 PVPCLQPNPQSQQOEQVPLVQ--QQQPPGQQ--QQPFPQPPQPPFPSPQPPYLQ 59  
DB 495 PLOEPQOEPPQOEPPQOEPPQOEPPQOEPPQOEPPQOEPPQOEPPQOEPPQOE 553  
QY 60 P---FPQ---PYLPYPQ---PQSFPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 109  
DB 554 PQRREPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 613  
QY 110 QQQQILQQLIQQLICMVDVLAQHNIARSGVLAQSTYQLLQLCCQHLMOIPESQC 169  
DB 614 EQQDDEQQDEQQ---QDEQQDEQQDEQQDEQQDEQQDEQQDE---QQDEQQD 663  
QY 170 QAIHNVHAILHQQQKQQQPPSSQVSPQPLQYPLQGGSPQPPQPPQPPQPPQPPQ 229  
DB 664 EQQD-----EQQDDEQQDEQQDEQQDEQQDEQQDEQQDEQQDEQQDEQQD 716  
QY 230 PQFEE 234  
DB 717 EQQD 721

RESULT 5  
US-09-894-273-2  
; Sequence 2, Application US/09894273  
; Patent No. 6756203  
; GENERAL INFORMATION:  
; APPLICANT: Kieff, Elliott D.  
; APPLICANT: Ballestas, Mary E.  
; APPLICANT: Kaye, Kenneth M.  
; TITLE OF INVENTION: RHADINO VIRUS LANA ACTS IN TRANS ON A UNIT OF RHADINO  
; FILE REFERENCE: 16412-10001R  
; CURRENT APPLICATION NUMBER: US/09/894,273  
; CURRENT FILING DATE: 2001-06-28  
; PRIOR APPLICATION NUMBER: US 60/109,422  
; PRIOR FILING DATE: 1998-11-19  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1162  
; TYPE: PRT  
; ORGANISM: Kaposi's sarcoma-associated herpesvirus  
US-09-894-273-2

Query Match 18.6%; Score 266; DB 4; Length 1162;  
Best Local Similarity 38.0%; Pred. No. 2.3e-16;  
Matches 93; Conservative 19; Mismatches 101; Indels 32; Gaps 10;

QY 4 PVPCLQPNPQSQQOEQVPLVQ--QQQPPGQQ--QQPFPQPPQPPFPSPQPPYLQ 59  
DB 495 PLOEPQOEPPQOEPPQOEPPQOEPPQOEPPQOEPPQOEPPQOEPPQOE 553  
QY 60 P---FPQ---PYLPYPQ---PQSFPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 109  
DB 554 PQRREPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 613  
QY 110 QQQQILQQLIQQLICMVDVLAQHNIARSGVLAQSTYQLLQLCCQHLMOIPESQC 169  
DB 614 EQQDDEQQDEQQ---QDEQQDEQQDEQQDEQQDEQQDEQQDE---QQDEQQD 663  
QY 170 QAIHNVHAILHQQQKQQQPPSSQVSPQPLQYPLQGGSPQPPQPPQPPQPPQPPQ 229  
DB 664 EQQD-----EQQDDEQQDEQQDEQQDEQQDEQQDEQQDEQQDEQQDEQQD 716  
QY 230 PQFEE 234  
DB 717 EQQD 721

DB 495 PLOEPQOEPPQOEPPQOEPPQOEPPQOEPPQOEPPQOEPPQOEPPQOE 553  
QY 60 P---FPQ---PYLPYPQ---PQSFPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 109  
DB 554 PQRREPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 613  
QY 110 QQQQILQQLIQQLICMVDVLAQHNIARSGVLAQSTYQLLQLCCQHLMOIPESQC 169  
DB 614 EQQDDEQQDEQQ---QDEQQDEQQDEQQDEQQDEQQDEQQDE---QQDEQQD 663  
QY 170 QAIHNVHAILHQQQKQQQPPSSQVSPQPLQYPLQGGSPQPPQPPQPPQPPQPPQ 229  
DB 664 EQQD-----EQQDDEQQDEQQDEQQDEQQDEQQDEQQDEQQDEQQDEQQD 716  
QY 230 PQFEE 234  
DB 717 EQQD 721

RESULT 6  
US-08-918-914-4  
; Sequence 4, Application US/08918914  
; Patent No. 5876963  
; GENERAL INFORMATION:  
; APPLICANT: Mitchell, Peter  
; APPLICANT: Hutchinson, Nancy  
; APPLICANT: Lawton, Michael  
; APPLICANT: Magana, Holly  
; APPLICANT: Yocum, Sue  
; APPLICANT: Murry, Lynn E.  
; TITLE OF INVENTION: HUMAN NUCLEOTIDE PYROPHOSPHORYLASE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/918,914  
FILING DATE: Filed Herewith  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0369  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-855-0555  
TELEFAX: 415-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 788 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: Genbank  
CLONE: 1070094  
US-08-918-914-4

Query Match 17.2%; Score 247; DB 2; Length 788;  
Best Local Similarity 31.0%; Pred. No. 8.3e-15;  
Matches 89; Conservative 23; Mismatches 97; Indels 78; Gaps 10;















Wed Dec 15 10:02:04 2004

us-10-089-700-3-y65.rai

Page 7

Search completed: December 14, 2004, 17:29:01  
Job time : 21 secs

---



*This Page Blank (uspto)*











Sequence 101, Application US/10474955  
 Publication No. US20040241161A1  
 GENERAL INFORMATION:  
 APPLICANT: Drifhout, Jan W.  
 APPLICANT: Konig, Frits  
 APPLICANT: McAdam, Stephen N.  
 APPLICANT: Ludwig, Solliid Magne  
 TITLE OF INVENTION: METHODS AND MEANS FOR USE OF HLA-DQ RESTRICTED T-CELL RECEPTORS A  
 FILE REFERENCE: 279/71244-PCT-US  
 CURRENT APPLICATION NUMBER: US/10/474,955  
 CURRENT FILING DATE: 2003-10-13  
 NUMBER OF SEQ ID NOS: 137  
 SOFTWARE: Patent in version 3.1  
 SEQ ID NO 101  
 LENGTH: 282  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Amino acid sequence of GAMMA-1  
 US-10-474-955-101

Query Match 38.8%; Score 556.5; DB 17; Length 282;  
 Best Local Similarity 47.6%; Pred. No. 2.7e-36;  
 Matches 138; Conservative 32; Mismatches 73; Indels 47; Gaps 15;

QY 2 RVVPFQL---OPNPSQQQPPQGVPLVQQQPPQ-QQF---PPQPPPPQ-QPPPS 51  
 DB 10 QVWFPQQQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 68  
 QY 52 --QQVYLQ--LQPPPPPPPP--YPPQPPPPPPPPPPPPPPPPPPPPPPPPPP 106  
 DB 69 PPQPP 118  
 QY 107 QQQQQQQQILQQLIPCMQDVVLQOHN---IAHRSQVLOQSTYQLQELCCQHLMQ 162  
 DB 119 PPQQQPSLQSLQQLQCNCKRFLQCKKPSVLSWSMLPRSDQVWQQCCQQLAQ 178  
 QY 163 IPESQCCAIHNVVHAIIHQCKQKQKQSSQVSTQCP--QQYPLGGGSRPSSQNPQAQ 221  
 DB 179 IPQQLQCAIHSHIVSHIQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQ 229  
 QY 222 GSVPPQQLPQFEIRNLALQTLPAQCNYIAPYCTI--APF-----GIFG 264  
 DB 230 GIIQPPQALVSHISLVLTATMCNVYVPPECSITAPFASVAGIGG 279

RESULT 6  
 US-10-739-930-9621  
 Sequence 9621, Application US/10739930  
 Publication No. US20040216190A1  
 GENERAL INFORMATION:  
 APPLICANT: Kowalik, David K.  
 TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH  
 FILE REFERENCE: 38-21(53377)B  
 CURRENT APPLICATION NUMBER: US/10/739,930  
 CURRENT FILING DATE: 2003-12-18  
 NUMBER OF SEQ ID NOS: 11088  
 SEQ ID NO 9621  
 LENGTH: 298  
 TYPE: PRT  
 ORGANISM: Trifidum aestivum  
 FEATURE:  
 OTHER INFORMATION: Clone ID: TRIAE-23APR03-C125\_65.P  
 US-10-739-930-9621

Query Match 38.1%; Score 545.5; DB 17; Length 298;  
 Best Local Similarity 46.9%; Pred. No. 2.2e-35;  
 Matches 134; Conservative 28; Mismatches 75; Indels 49; Gaps 14;

QY 4 PVPQQLPQNPSSQQQPPQGVPLVQQQPPQPPQPPQPPQPPQPPQPPQPPQPP 61

DB 36 PVP--QHPQPFSSQP-----QTFPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 84  
 QY 62 P-QPVLPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 110  
 DB 85 PQQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 140  
 QY 111 QQQILQQLIQQQLIPCMQDVVLQOHN---IAHRSQVLOQSTYQLQELCCQHLMQ 166  
 DB 141 QPPPIQPSLQQQVNPCKNFLQCKKPSVLSWSMLWPSQCVWQQCCQQLAQIPQ 200  
 QY 167 SQCCAIHNVVHAIIHQCKQKQKQSSQVSTQCP--QQYPLGGGSRPSSQNPQAQSVQ 225  
 DB 201 LQCAIHHTIHSITMQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQ 251  
 QY 226 PQQLPQFEIRNLALQTLPAQCNYIAPYCTI--APF-----GIFG 264  
 DB 252 PQQPQALVSHISLVLTATMCNVYVPPECSITAPFASVAGIGG 297

RESULT 7  
 US-10-474-955-97  
 Sequence 97, Application US/10474955  
 Publication No. US20040241161A1  
 GENERAL INFORMATION:  
 APPLICANT: Drifhout, Jan W.  
 APPLICANT: Konig, Frits  
 APPLICANT: McAdam, Stephen N.  
 APPLICANT: Ludwig, Solliid Magne  
 TITLE OF INVENTION: METHODS AND MEANS FOR USE OF HLA-DQ RESTRICTED T-CELL RECEPTORS  
 FILE REFERENCE: 279/71244-PCT-US  
 CURRENT APPLICATION NUMBER: US/10/474,955  
 CURRENT FILING DATE: 2003-10-13  
 NUMBER OF SEQ ID NOS: 137  
 SOFTWARE: Patent in version 3.1  
 SEQ ID NO 97  
 LENGTH: 279  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Consensus amino acid sequence  
 US-10-474-955-97

Query Match 36.5%; Score 523; DB 17; Length 279;  
 Best Local Similarity 46.2%; Pred. No. 1.2e-33;  
 Matches 132; Conservative 26; Mismatches 76; Indels 52; Gaps 14;

QY 4 PVPQQLPQNPSSQQQPPQGVPLVQQQPPQPPQPPQPPQPPQPPQPPQPPQPP 61  
 DB 18 PVP--QHPQPFSSQP-----QTFPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 66  
 QY 62 P-QPVLPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 110  
 DB 67 PQQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPP 122  
 QY 111 QQQILQQLIQQQLIPCMQDVVLQOHN---IAHRSQVLOQSTYQLQELCCQHLMQ 166  
 DB 123 QPPPIQPSLQQQVNPCKNFLQCKKPSVLSWSMLWPSQCVWQQCCQQLAQIPQ 182  
 QY 167 SQCCAIHNVVHAIIHQCKQKQKQSSQVSTQCP--QQYPLGGGSRPSSQNPQAQSVQ 225  
 DB 183 LQCAIHHTIHSITMQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQEQ 230  
 QY 226 PQQLPQFEIRNLALQTLPAQCNYIAPYCTI--APF-----GIFG 264  
 DB 231 PQQPQALVSHISLVLTATMCNVYVPPECSITAPFASVAGIGG 276

RESULT 8  
 US-10-474-955-99  
 Sequence 99, Application US/10474955  
 Publication No. US20040241161A1  
 GENERAL INFORMATION:



APPLICANT: Drifhout, Jan W.  
 APPLICANT: Konig, Frits  
 APPLICANT: McAdam, Stephan N.  
 APPLICANT: Ludwig, Solid Magne  
 TITLE OF INVENTION: METHODS AND MEANS FOR USE OF HLA-DO RESTRICTED T-CELL RECEPTORS A  
 FILE REFERENCE: 2799/71244-PCT-US  
 CURRENT APPLICATION NUMBER: US/10/474,955  
 CURRENT FILING DATE: 2003-10-13  
 NUMBER OF SEQ ID NOS: 137  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 99  
 LENGTH: 279  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Amino acid sequence of GAMMA-4  
 US-10-474-955-99  
 Query Match  
 Best Local Similarity 36.2%; Score 522; DB 17; Length 279;  
 Matches 132; Conservative 26; Mismatches 76; Indels 52; Gaps 14;  
 DB 4 PVPQLQPNPESQQQPQSEVPLVQQQQPFGQQQQPF--PQQPYQPQPFPSQQPYLQLQPF 61  
 18 PVP--QHPQPSQP-----QQTFFPQQTFFHQPPQQQFPQPPQ--PQQQLQPPQPF 66  
 62 P-QPYLPPYPPQ--PQSFPPQPPYPPQ-----PQYSQPPQPPISQQQAQ--QQQQQQQQQQ 110  
 67 PQQPQQPPYPPQ--QPPQPPQQTQPPQLPQSQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 122  
 111 QQQILQQLLQQQLPCHMDVYLQGN-----IAHARSQVLQSTYQLLOELCCGHLMOIPQ 166  
 123 QPPIQPSLQQQVNPCKNFFLLQCKPVSLSWSMTWPSDDCVMRQSCQQLAQLPQ 182  
 167 SOCCAHHNVVAIILHQQKQQQPPSSQVSFQPL--QPYLQGSFPPSQNPQAQGSVP 225  
 183 LQCAIHHTVHSITIMQEQQQ-----GMHLLPLVYQQQVGGQTL-----VQGGGIIQ 230  
 226 PQQLPQFEIRNALQTLPMKCNVYIAPYCTI--APF-----GTFG 264  
 231 PQQPQLFAIRSLVLTLPKCNVYVPECSITIKAPSSVAVAGIG 276  
 RESULT 9  
 US-10-474-955-98  
 Sequence 98, Application US/10474955  
 Publication No. US20040241161A1  
 GENERAL INFORMATION:  
 APPLICANT: Drifhout, Jan W.  
 APPLICANT: Konig, Frits  
 APPLICANT: McAdam, Stephan N.  
 APPLICANT: Ludwig, Solid Magne  
 TITLE OF INVENTION: METHODS AND MEANS FOR USE OF HLA-DO RESTRICTED T-CELL RECEPTORS A  
 FILE REFERENCE: 2799/71244-PCT-US  
 CURRENT APPLICATION NUMBER: US/10/474,955  
 CURRENT FILING DATE: 2003-10-13  
 NUMBER OF SEQ ID NOS: 137  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 98  
 LENGTH: 279  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Amino acid sequence of GAMMA-2  
 US-10-474-955-98  
 Query Match  
 Best Local Similarity 36.2%; Score 519; DB 17; Length 279;  
 Matches 131; Conservative 26; Mismatches 77; Indels 52; Gaps 14;  
 DB 4 PVPQLQPNPESQQQPQSEVPLVQQQQPFGQQQQPF--PQQPYQPQPFPSQQPYLQLQPF 61  
 18 PVP--QHPQPSQP-----QQTFFPQQTFFHQPPQQQFPQPPQ--PQQQLQPPQPF 66  
 62 P-QPYLPPYPPQ--PQSFPPQPPYPPQ-----PQYSQPPQPPISQQQAQ--QQQQQQQQQQ 110  
 67 PQQPQQPPYPPQ--QPPQPPQQTQPPQLPQSQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 122  
 111 QQQILQQLLQQQLPCHMDVYLQGN-----IAHARSQVLQSTYQLLOELCCGHLMOIPQ 166  
 123 QPPIQPSLQQQVNPCKNFFLLQCKPVSLSWSMTWPSDDCVMRQSCQQLAQLPQ 182  
 167 SOCCAHHNVVAIILHQQKQQQPPSSQVSFQPL--QPYLQGSFPPSQNPQAQGSVP 225  
 183 LQCAIHHTVHSITIMQEQQQ-----GMHLLPLVYQQQVGGQTL-----VQGGGIIQ 230  
 226 PQQLPQFEIRNALQTLPMKCNVYIAPYCTI--APF-----GTFG 264  
 231 PQQPQLFAIRSLVLTLPKCNVYVPECSITIKAPSSVAVAGIG 276

DB 18 PVP--QHPQPSQP-----QQTFFPQQTFFHQPPQQQFPQPPQ--PQQQLQPPQPF 66  
 QY 62 P-QPYLPPYPPQ--PQSFPPQPPYPPQ-----PQYSQPPQPPISQQQAQ--QQQQQQQQQQ 110  
 DB 67 PQQPQQPPYPPQ--QPPQPPQQTQPPQLPQSQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 122  
 QY 111 QQQILQQLLQQQLPCHMDVYLQGN-----IAHARSQVLQSTYQLLOELCCGHLMOIPQ 166  
 DB 123 QPPIQPSLQQQVNPCKNFFLLQCKPVSLSWSMTWPSDDCVMRQSCQQLAQLPQ 182  
 QY 167 SOCCAHHNVVAIILHQQKQQQPPSSQVSFQPL--QPYLQGSFPPSQNPQAQGSVP 225  
 DB 183 LQCAIHHTVHSITIMQEQQQ-----GMHLLPLVYQQQVGGQTL-----VQGGGIIQ 230  
 QY 226 PQQLPQFEIRNALQTLPMKCNVYIAPYCTI--APF-----GTFG 264  
 DB 231 PQQPQLFAIRSLVLTLPKCNVYVPECSITIKAPSSVAVAGIG 276  
 RESULT 10  
 US-10-474-955-100  
 Sequence 100, Application US/10474955  
 Publication No. US20040241161A1  
 GENERAL INFORMATION:  
 APPLICANT: Drifhout, Jan W.  
 APPLICANT: Konig, Frits  
 APPLICANT: McAdam, Stephan N.  
 APPLICANT: Ludwig, Solid Magne  
 TITLE OF INVENTION: METHODS AND MEANS FOR USE OF HLA-DO RESTRICTED T-CELL RECEPTORS  
 FILE REFERENCE: 2799/71244-PCT-US  
 CURRENT APPLICATION NUMBER: US/10/474,955  
 CURRENT FILING DATE: 2003-10-13  
 NUMBER OF SEQ ID NOS: 137  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 100  
 LENGTH: 279  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Amino acid sequence of GAMMA-3  
 US-10-474-955-100  
 Query Match  
 Best Local Similarity 36.2%; Score 519; DB 17; Length 279;  
 Matches 131; Conservative 28; Mismatches 76; Indels 50; Gaps 14;  
 DB 4 PVPQLQPNPESQQQPQSEVPLVQQQQPFGQQQQPF--PQQPYQPQPFPSQQPYLQLQPF 61  
 18 PVP--QHPQPSQP-----QQTFFPQQTFFHQPPQQQFPQPPQ--PQQQLQPPQPF 66  
 62 P-QPYLPPYPPQ--PQSFPPQPPYPPQ-----PQYSQPPQPPISQQQAQ--QQQQQQQQQQ 111  
 67 PQQPQQPPYPPQ--QPPQPPQQTQPPQLPQSQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 123  
 QY 112 QQQILQQLLQQQLPCHMDVYLQGN-----IAHARSQVLQSTYQLLOELCCGHLMOIPQ 167  
 DB 124 PPIQPSLQQQVNPCKNFFLLQCKPVSLSWSMTWPSDDCVMRQSCQQLAQLPQ 183  
 QY 168 SOCCAHHNVVAIILHQQKQQQPPSSQVSFQPL--QPYLQGSFPPSQNPQAQGSVP 226  
 DB 184 LQCAIHHTVHSITIMQEQQQ-----GMHLLPLVYQQQVGGQTL-----VQGGGIIQ 231  
 QY 227 PQQLPQFEIRNALQTLPMKCNVYIAPYCTI--APF-----GTFG 264  
 DB 232 PQQPQLFAIRSLVLTLPKCNVYVPECSITIKAPSSVAVAGIG 276  
 RESULT 11  
 US-10-739-930-9778  
 Sequence 9778, Application US/10739930  
 Publication No. US2004021619CA1



Query Match	30.6%	Score 438.5	DB 17	Length 307
Best Local Similarity	42.8%	Ped. No. 7.1e-27		
Matches 125	Conservative 36	Mismatches 76	Indels 55	Gaps 16

  

QY	13	BSQQQPPQGVPLVQQQQQPF-----GQQQQPPQQYYPQGGPPSPSQPPYQLQAPFPQPIPL	67
DB	27	PELEBPMQQQQLPMPQQPFPQQPLFSQQQQ--QQLEPQQPSFSQQPF-----PFPQQQPF	78
QY	68	YFPQPSFPQGP-----YFPQPPQYSGQPPFI--SQQAQQQQQQQQQQQQQQQI	114
DB	79	FSQQQPIPLPQQPPFSQQQQQLVLPQ--QPPSSQQQQLVLPQSGPPFPQQQQQQHQQLVQQQLP	137

```

RESULT 14
US-10-425-115-200100
; Sequence 200100, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
;

```















[illegible]

RESULT 6  
S07361  
alpha/beta-gliadin precursor (clone pW1215) - wheat  
C/Species: Triticum aestivum (common wheat)  
C/Date: 08-Jun-1994 #sequence\_revision 01-Dec-1995 #text\_change 09-Jul-2004  
C/Accession: S07361  
R/Submitter-Smith, M.; Rafalski, J.A.; Sugiyama, T.; Stoll, M.; Soell, D.  
Nucleic Acids Res. 13, 3905-3916, 1985  
A/Title: Conservation and variability of wheat alpha/beta-gliadin genes.  
A/Reference number: S07361; PMID:85242077; PMID:3839304  
A/Accession: S07361  
A/Status: Preliminary  
A/Molecule type: DNA  
A/Residues: 1-296 <SUM>  
A/Cross-references: UNIPROT:P04726; EMBL:X02538; NID:g21756; PIDN:CAA26383.1; PID:g21755  
C/Superfamily: gliadin  
C/Keywords: seed; storage protein

Query Match 88.0%; Score 1261; DB 2; Length 296;  
Best Local Similarity 88.1%; Pred. No. 1.7e-77;  
Matches 244; Conservative 4; Mismatches 17; Indels 12; Gaps 4;

QY	1	VAVPVPOLQPNPSSQOQPEQVPLVQOQOQFPQOQOQFPQOQPYQCPQPPSQOPLYLOLP	60
Db	21	VAVPVPQPNPQSPQOQOQVPLVQOQOQFPQOQOQFPQOQPYQCPQPPSQOPLYLOLP	80
QY	61	FPQ-----PYLEPQOQSFPPQOQPYQOQPYQOQPISSQOQ-----QQOQOQOQOQOQOQ	111
Db	81	FPQOQFPPLQLYPQCPFPSPQOQPYQOQPYQOQCPQPISSQOQOQOQOQOQOQOQOQOQOQ	140
QY	112	QOILQOILQOQILPCQDQVVLQOQHNINHARSQVLYQSTYOLLBELCCQHLMOJPEQSOQCA	171
Db	141	QOILQOILQOQILPCQDQVVLQOQHNINHARSQVLYQSTYOLLBELCCQHLMOJPEQSRCA	200
QY	172	HNHVAHATLHQOQKQOQSSQVSFOQLQOQPYLQGSFRRPQOQNPQAGSVQPOQLPQ	231
Db	201	HNHVAHATLHQOQR-QOQPSQVSLQOQOQOQPSQOQGFQFQOQNPQAGSVQPOQLPQ	259
QY	232	FEIRINLATQTLPAWGNVTIAPC--TIAPEGPGTN	266
Db	260	FEIRINLATQTLPRMGNVTIAPCSTTIAPEGPGTN	296

```

RESULT 7
S07924
alpha/beta-gliadin precursor - wheat
C1:Species: Triticum aestivum (common wheat)
C1:date: 08-Jun-1994 #revision_01-Dec-1995 #next_change 09-Jul-2004
C1:accession: S07924; C61218

```

R:Summer-Smith, M.; Rafalski, J.A.; Sugiyama, T.; Stoll, M.; Scell, D.  
Nucleic Acids Res. 13, 3905-3916, 1985  
A>Title: Conservation and variability of wheat alpha/beta-gliadin genes.  
A:Reference number: S07361, MUID:65242077, PMID:383304  
A:Accession: S07924  
A>Status: preliminary; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-313 <STU>  
A:Cross-references: UNIPROT:Q41546; EMBL:X02540; NID:G21764; PID:G2176  
R:Shewry, P.R.; Sabelli, P.A.; Parmar, S.; Lafindra, D.  
Biochem. Genet. 29, 207-211, 1991  
A>Title: alpha-type prolamins are encoded by genes on chromosomes 4Ha and 6Ha of Haynald  
A:Reference number: A61218; MUID:91315394; PMID:1855356  
A:Accession: C61218  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 18-27 <SH>  
C:Superfamily: gliadin  
C:Keywords: seed; storage protein

[illegible]

RESULT 8  
 A27319  
 gliadin - wheat  
 C:Species: Triticum aestivum (common wheat)  
 C:Date: 04-Mar-1988 #sequence\_revision 04-Mar-1988 #text\_change 03-Feb-1994  
 C:Accession: A27319  
 R:Reeves, C.D.; Okita, T.W.  
 Gene 52, 257-266, 1987  
 A:Title: Analyses of alpha/beta-type gliadin genes from diploid and hexaploid wheats  
 A:Reference number: A27319, MUID:87277398, PMID:3038689  
 A:Accession: A27319  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-296 <RES>  
 C:Superfamily: gliadin

Query Match	87.6%	Score 1255	DB 2	Length 236	
Best Local Similarity	87.7%	Pred. No. 4	4e-77		
Matches 243	Conservative	3	Mismatches 19	Incls 12	Gaps 4

[illegible]











A:Residues: 1-292 <OKI>  
A:Cross-references: UNIPROT:P04721  
E:Superfamily: gliadin  
F:1-20/Domain: signal sequence #status predicted <SIG>  
F:21-292/Product: alpha/beta-gliadin #status predicted <MAT>

Query Match 80.8%; Score 1157.5; DB 2; length 292;  
Best Local Similarity 84.1%; Pred. No. 1,4e-70;  
Matches 227; Conservative 3; Mismatches 9; Indels 31; Gaps 4;

```
DB 1 VRVVPQLQPNPSSQQLPQECVPLVQQQCPFGQQQFPFQQQYPQDPFPSSQGPYLQLQP 60
    |||||
DB 21 VRVVPQLQGNPQQQPPQGVPLVQQQFLGQQPPFPQQYPQDPFPSSQGPYLQLQP 80
    |||||
DB 61 F--PQPY---LPYPQSSFPQQPYPQDPQYSQPQPIS--QQQAQQQQQQQQQQQQQ 114
    |||||
DB 81 FLQPQPPPPQLPYSPQPPFPQQPQYSPQQPISQQQQQQQQQQQQQQQQQQQ 140
    |||||
DB 115 LQQILQQQLPCMDVVLQGNHIAHARSQVLQGSFYQLLQELCCQHLWQIPQSSQCAIHN 174
    |||||
DB 141 IQQLLQQQLPCMDVVLQGNHIVHGSQVLQGSFYQLLQELCCQHLWQIPQSSQCAIHN 200
    |||||
DB 175 VVAHIIHH-----QQKQQQSSQVVSFQQPLQQYPLGGG 209
    |||||
DB 201 VVAHIIHHQQQQQQEQKQLQQQQQQQQQLQQQQQQKQQQQSSQVVSFQQPLQQYPLGGG 260
    |||||
DB 210 -SFRPSSQNPQKQSSVQPPQLPQFEIRNLA 239
    |||||
DB 261 SFRPSSQNPQKQSSVQPPQLPQFEIRNLA 290
    |||||
```

Search completed: December 14, 2004, 17:26:59  
Job time : 15 secs











## RESULT 4

Q9ZP09 PRELIMINARY; PRT; 288 AA.

ID Q9ZP09; 01-MAY-1999 (TREMBlrel. 10, Created)

AC Q9ZP09; 01-MAY-1999 (TREMBlrel. 10, Last sequence update)

DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)

DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)

DE Alpha-gliadin precursor (Fragment).

GN Name=alpha-gliadin;

OS Triticum aestivum subsp. spelta.

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;

OC Triticaceae; Triticum.

NCBI\_TaxID=58933;

RN [1]

RP SEQUENCE FROM N.A.

RA Kaaard D.D.; D'Ovidio R.;

RT "Deduced amino acid sequences of an alpha-gliadin gene from Spelt wheat

RL Cereal Chem. 76:548-551(1999).

DR EMBL; AJ10948; CAI0257.1; -.

DR PIR; G13333; S13333.

DR GO; GO:0045735; F:nutrient reservoir activity; IEA.

DR InterPro; IPR003612; AAI.

DR InterPro; IPR001376; Gliadin.

DR InterPro; IPR001954; Gli\_a\_glutenin.

DR Pfam; PF00234; TYP\_alpha\_amy1; 1.

DR PRINTS; PR00208; GLIADGLUTEN.

DR PRINTS; PR00209; GLIADIN.

DR SMART; SM00499; AAI; 1.

DR KMW SIGNAL.

FT CHAIN 1 20 Potential.

FT NON TER 21 >288 alpha-gliadin.

FT 288 288

SEQUENCE 288 AA; 33203 MW; DA058F3FAFA6BC6C CRC64;

Query Match 94.3%; Score 1352; DB 2; Length 288;

Best Local Similarity 95.1%; Pred. No. 5,4e-75;

Matches 255; Conservative 0; Mismatches 11; Indels 2; Gaps 1;

QY 1 VRVVPQLQPNPSSQOQPOEQVPLVQOQOFPQOQOQFPPOQPYPOPPFPSSQOQPYLOQ 60

DB 21 VRVVPQLQPNPSSQOQPOEQVPLVQOQOFLGQOQFPFPQPYPOPPFPSSQOQPYLOQ 80

QY 61 FPQPYLPYPOPSFPPOQPYPOPPQYSPQPPIS--QQQAQQOQQOQQOQQOQQOQ 118

DB 81 FPQPYLPYPOPSFPPOQPYPOPPQYSPQPPISQQOQQOQQOQQOQQOQQOQ 140

QY 119 LQQOQLPCMDVYLQCHNIAHARSQVLAQSTYQLQLGLCCOHLMQIPESQCCOAIHNV 178

DB 141 LQQOQLPCMDVYLQCHNIAHARSQVLAQSTYQLQLGLCCOHLMQIPESQCCOAIHNV 200

QY 179 IILHQOQKQOQOQSSQVSPQPLQOQYPLGQSFRRSQNPQAQGSVQPOQLPQFEI 238

DB 201 IILHQOQKQOQOQSSQVSPQPLQOQYPLGQSFRRSQNPQAQGSVQPOQLPQFEI 260

QY 239 ALQTLPMACNVYIAPYCTIAPFGIRGTN 266

DB 261 ALQTLPMACNVYIAPYCTIAPFGIRGTN 288

RESULT 5

Q9M4M5 PRELIMINARY; PRT; 274 AA.

ID Q9M4M5; 01-OCT-2000 (TREMBlrel. 15, Created)

AC Q9M4M5; 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)

DE Alpha-gliadin

OS Triticum aestivum (wheat).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;

OC Triticaceae; Triticum.

OX NCBI\_TaxID=4565;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Mjoelner; TISSUE=Endosperm;

RA Arentz-Hansen E.H.; McAdam S.N.; Molberg O.; Kristiansen C.;

RA Sollid L.M.;

RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ133603; CAB76955.1; -.

DR GO; GO:0045735; F:nutrient reservoir activity; IEA.

DR InterPro; IPR003612; AAI.

DR InterPro; IPR001376; Gliadin.

DR InterPro; IPR001954; Gli\_a\_glutenin.

DR Pfam; PF00234; TYP\_alpha\_amy1; 1.

DR PRINTS; PR00208; GLIADGLUTEN.

DR PRINTS; PR00209; GLIADIN.

DR SMART; SM00499; AAI; 1.

FT CHAIN 1 274

SEQUENCE 274 AA; 31980 MW; 976919397534ABD CRC64;

Query Match 94.1%; Score 1348.5; DB 2; Length 274;

Best Local Similarity 94.1%; Pred. No. 8,5e-75;

Matches 255; Conservative 0; Mismatches 11; Indels 5; Gaps 1;

QY 1 VRVVPQLQPNPSSQOQPOEQVPLVQOQOFPQOQOQFPPOQPYPOPPFPSSQOQPYLOQ 60

DB 2 VRVVPQLQPNPSSQOQPOEQVPLVQOQOFLGQOQFPFPQPYPOPPFPSSQOQPYLOQ 61

QY 61 FPQPYLPYPOPSFPPOQPYPOPPQYSPQPPIS----QQQAQQOQQOQQOQQOQ 115

DB 62 FPQPYLPYPOPSFPPOQPYPOPPQYSPQPPISQQOQQOQQOQQOQQOQQOQ 121

QY 116 QQLLQOQLPCMDVYLQCHNIAHARSQVLAQSTYQLQLGLCCOHLMQIPESQCCOAIHNV 175

DB 122 QQLLQOQLPCMDVYLQCHNIAHARSQVLAQSTYQLQLGLCCOHLMQIPESQCCOAIHNV 181

QY 176 VHAIIHQOQKQOQOQSSQVSPQPLQOQYPLGQSFRRSQNPQAQGSVQPOQLPQFEI 235

DB 182 VHAIIHQOQKQOQOQSSQVSPQPLQOQYPLGQSFRRSQNPQAQGSVQPOQLPQFEI 241

QY 236 RNIALQTLPMACNVYIAPYCTIAPFGIRGTN 266

DB 242 RNIALQTLPMACNVYIAPYCTIAPFGIRGTN 272

RESULT 6

Q9M4M2 PRELIMINARY; PRT; 276 AA.

ID Q9M4M2; 01-OCT-2000 (TREMBlrel. 15, Created)

AC Q9M4M2; 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)

DE Alpha-gliadin

OS Triticum aestivum (wheat).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;

OC Triticaceae; Triticum.

NCBI\_TaxID=4565;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Mjoelner; TISSUE=Endosperm;

RA Arentz-Hansen E.H.; McAdam S.N.; Molberg O.; Kristiansen C.;

RA Sollid L.M.;

RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ133606; CAB76958.1; -.

DR GO; GO:0045735; F:nutrient reservoir activity; IEA.

DR InterPro; IPR003612; AAI.

DR InterPro; IPR001376; Gliadin.

DR InterPro; IPR001954; Gli\_a\_glutenin.

DR Pfam; PF00234; TYP\_alpha\_amy1; 1.

DR PRINTS; PR00208; GLIADGLUTEN.

DR PRINTS; PR00209; GLIADIN.

DR SMART; SM00499; AAI; 1.

FT CHAIN 1 276

alpha-gliadin.







[illegible][illegible]







DR EMBL; U51304; AAA96523.1; -.  
DR PIR; S13333; S13333.  
DR PIR; T06282; T06282.  
DR GO; GO:0045735; F:multient reservoir activity; IEA.  
DR InterPro; IPR003512; AAI.  
DR InterPro; IPR001376; Gliadin.  
DR InterPro; IPR001954; G1ia\_glutenin.  
DR Pfam; PF00234; TRY\_alpha\_aml; 1.  
DR PRINTS; PR00208; GLIADGLUTEN.  
DR PRINTS; PR00209; GLIADIN.  
DR SMART; SM00499; AAI; 1.  
SQ SEQUENCE 288 AA; 33285 MW; D7F6B99133283CA2 CRC64;

Query Match 91.4%; Score 1310; DB 2; Length 288;  
Best Local Similarity 92.5%; Pred. No. 2e-72;  
Matches 248; Conservative 3; Mismatches 15; Indels 2; Gaps 1;

QY 1 VRVPVQLQPNPNSQQLPQEQVPLVQQQQLGQQQFPFPQCPYPQPFPSQCPYLQLP 60  
DB 21 VRVPVQLQPNPNSQQLPQEQVPLVQQQQLGQQQFPFPQCPYPQPFPSQCPYLQLP 80  
QY 61 FPQPIPLPYPQPSFPFPQCPYPQPFPSQCPYPQPFPSQCPYLQLP 118  
DB 81 FPQPIPLPYPQPSFPFPQCPYPQPFPSQCPYPQPFPSQCPYLQLP 140  
QY 119 LQQQLIPCMQDVVLQCHNIAHARSQVLCQSTYQLLQELCCQHLMOIPEQSCQCAIHNVVHA 178  
DB 141 LQQQLIPCMQDVVLQCHNIAHARSQVLCQSTYQLLQELCCQHLMOIPEQSCQCAIHNVVHA 200  
QY 179 ILHQQK-QQQQPSQVSFQCPPLQCYPLQGSFRPSQNPQAGSVQPOLPQFEIRNL 238  
DB 201 ILHQQK-QQQQPSQVSFQCPPLQCYPLQGSFRPSQNPQAGSVQPOLPQFEIRNL 260  
QY 239 ALQTLPMQCNVYIAPYCTIAPFGIFGTN 266  
DB 261 ALQTLPMQCNVYIAPYCTIAPFGIFGTN 288

RESULT 15

Q41528 PRELIMINARY; PRT; 287 AA.

AC Q41528;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
DE Alpha-gliadin.  
OS Triticum aestivum (wheat).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
OC Triticeae; Triticum.  
OX NCBI\_TaxID=4565;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Cheyenne;  
RA Anderson O.D.;  
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U50984; AAA96276.1; -.  
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.  
DR InterPro; IPR003512; AAI.  
DR InterPro; IPR001376; Gliadin.  
DR InterPro; IPR001954; G1ia\_glutenin.  
DR Pfam; PF00234; TRY\_alpha\_aml; 1.  
DR PRINTS; PR00208; GLIADGLUTEN.  
DR PRINTS; PR00209; GLIADIN.  
DR SMART; SM00499; AAI; 1.  
SQ SEQUENCE 287 AA; 33193 MW; 05F82296749CE97 CRC64;

Query Match 90.5%; Score 1297.5; DB 2; Length 287;  
Best Local Similarity 92.9%; Pred. No. 1.1e-71;  
Matches 249; Conservative 2; Mismatches 14; Indels 3; Gaps 3;

QY 1 VRVPVQLQPNPNSQQLPQEQVPLVQQQQLGQQQFPFPQCPYPQPFPSQCPYLQLP 60

DB 21 VRVPVQLQPNPNSQQLPQEQVPLVQQQQLGQQQFPFPQCPYPQPFPSQCPYLQLP 79  
QY 61 FPQPIPLPYPQPSFPFPQCPYPQPFPSQCPYPQPFPSQCPYLQLP 119  
DB 80 FPQPIPLPYPQPSFPFPQCPYPQPFPSQCPYPQPFPSQCPYLQLP 139  
QY 120 LQQQLIPCMQDVVLQCHNIAHARSQVLCQSTYQLLQELCCQHLMOIPEQSCQCAIHNVVHA 179  
DB 140 LQQQLIPCMQDVVLQCHNIAHARSQVLCQSTYQLLQELCCQHLMOIPEQSCQCAIHNVVHA 199  
QY 180 ILHQQK-QQQQPSQVSFQCPPLQCYPLQGSFRPSQNPQAGSVQPOLPQFEIRNL 238  
DB 200 ILHQQK-QQQQPSQVSFQCPPLQCYPLQGSFRPSQNPQAGSVQPOLPQFEIRNL 259  
QY 239 ALQTLPMQCNVYIAPYCTIAPFGIFGTN 266  
DB 260 ALQTLPMQCNVYIAPYCTIAPFGIFGTN 287

Search completed: December 14, 2004, 17:25:22  
Job time : 80 secs



This Page Blank (uspto)